Rosenblatt, M., *Remarks on a multivariate transformation*, Ann. Math. Statist. **23** (1952), 470–472.

Sakhanenko, A. I., *Rate of convergence in the invariance principles for variables with exponential moments that are not identically distributed*, In: Trudy Inst. Mat. SO AN SSSR, vol. 3, Nauka, Novosibirsk, 1984, pp. 4–49. (Russian)

Zaitsev, A. Yu., *Estimates of the Lévy–Prokhorov distance in the multivariate central limit theorem for random variables with finite exponential moments*, Theor. Probab. Appl. **31** (1986), no. 2, 203–220.

Zaitsev, A. Yu., *Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments.*, Preprint 95-055 SFB 343, Universität Bielefeld (1995).

Zaitsev, A. Yu., *Estimates for quantiles of smooth conditional distributions and multidimensional invariance principle*, Siberian Math. J. **37** (1996a), no. 4, 807–831. (Russian)

Zaitsev, A. Yu., *Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments*, submitted to ESAIM : Probability and Statistics (1996b).

Friedrich Götze

Fakultät für Mathematik

Universität Bielefeld

Postfach 100131

33501 Bielefeld 1

Germany

E-mail address: goetze@mathematik.uni-bielefeld.de

Andrei Yu. Zaitsev

St. Petersburg Branch of Steklov Mathematical Institute

Fontanka 27

St. Petersburg 191011

Russia

E-mail address: zaitsev@pdmi.ras.ru

Put

$$c_{25} = \frac{(c_3 \log 2 + 1)}{c_2}, \qquad c_{26} = c_{25} \sum_{l=0}^{\infty} 2^{-l/2} = \frac{c_{25}\sqrt{2}}{\sqrt{2}-1}, \tag{46}$$

and introduce the events

$$A_l = \Big\{ \omega : \Delta^{(l)} \geqslant 2^l \, c_{26} \, \tau \, d^{3/2} \log^+ d \Big\}, \qquad l = 1, 2, \ldots, \tag{47}$$

where

$$\Delta^{(l)} = \max_{1 \leqslant r \leqslant m_l} \Big| \sum_{j=1}^{r} X_j - \sum_{j=1}^{r} Y_j \Big|. \tag{48}$$

According to (44), (45) and (48), we have

$$\Delta^{(l)} \leqslant \Delta_1 + \cdots + \Delta_l. \tag{49}$$

Taking into account the relations (42), (46), (47), (49) and applying the inequality (43) with $x = 2^{(s+l)/2}$, we get

$$\begin{aligned} \mathbf{P}\left\{ A_l \right\} &\leqslant \sum_{s=1}^{l} \mathbf{P}\Big\{ \Delta_s \geqslant 2^{(s+l)/2} \, c_{25} \, \tau \, d^{3/2} \log^+ d \Big\} \\ &\leqslant \sum_{s=1}^{l} \exp\big(-2^{(s+l)/2}\big) \leqslant c \, \exp\big(-2^{l/2}\big). \end{aligned} \tag{50}$$

The inequality (50) implies that $\sum\limits_{l=1}^{\infty} \mathbf{P}\left\{ A_l \right\} < \infty$, Hence, by the Borel–Cantelli lemma with probability one a finite number of the events A_l occurs only. This implies the equality (1.13) with $c_4 = 2\,c_{26} \big/ \log 2$ (see (41), (47) and (48)). □

References

Bártfai, P., *Die Bestimmung der zu einem wiederkehrenden Prozess gehörenden Verteilungfunktion aus den mit Fehlern behafteten Daten einer einzigen Realisation*, Studia Sci. Math. Hungar. **1** (1966), 161–168.

Dudley, R. M., *Real analysis and probability*, Wadsworth & Brooks/Cole, Pacific Grove, California, 1989.

Einmahl, U., *Extensions of results of Komlós, Major and Tusnády to the multivariate case*, J. Multivar. Anal. **28** (1989), 20–68.

Etemadi, N., *On some classical results in probability theory*, Sankhyā, Ser. A **47** (1985), no. 2, 215–221.

Hoffmann-Jørgensen, J., *Probability with a view toward statistics*, vol. I, Chapman & Hall, New York, 1994.

Komlós, J., Major, P., Tusnády, G., *An approximation of partial sums of independent RV'-s and the sample DF.* I; II, Z. Wahrscheinlichkeitstheor. verw. Geb. **32** (1975), 111–131; **34** (1976), 34–58.

Major, P., *On the invariance principle for sums of independent identically distributed random variables*, J. Multivar. Anal. **8** (1978), 487–517.

Massart, P., *Strong approximation for multivariate empirical and related processes, via KMT construction*, Ann. Probab. **17** (1989), no. 1, 266–291.

(see (1.10) and (27)). Below (in the proof of (5)) we assume that $\|h\|\sqrt{2}\,\tau < 1$. According to (37) and (38), for $t \in T$ we have

$$\bigl|\widehat{F}_h(t)\bigr|^{1/2} \leqslant e^{-m b_{10}\cdot 2^{n-2}}. \tag{39}$$

Taking into account that $\bigl|\widehat{F}_h(t)\bigr| \leqslant 1$, and $m \geqslant b_4$, choosing b_4 to be sufficiently large and using (1.10), (7), (36) and (39), we obtain

$$\begin{aligned}\int\limits_T \bigl|\widehat{F}_h(t)\bigr|\,dt &\leqslant \exp\bigl(-m b_{10}\cdot 2^{n-2}\bigr)\biggl(\int\limits_{\mathbb{R}^j}\Bigl(1+\frac{\|t\|}{b_7\sqrt{m}}\Bigr)^{-m\cdot 2^{n-2}} dt + b_{11}\,m^{d/2}\biggr)\\ &\leqslant b_{12}\,m^{d/2}\,\exp\bigl(-m b_{10}\cdot 2^{n-2}\bigr)\\ &\leqslant \frac{(2\,\pi)^{j/2}\sqrt{2}\,b_2\,j^{3/2}}{m^{1/2}\cdot 2^{n/2}\cdot 2^{nj/2}} = \frac{(2\,\pi)^{j/2}\,\sqrt{2}\,\tau\,j^{3/2}}{\sigma\,(\det\mathbb{D})^{1/2}}.\end{aligned} \tag{40}$$

The inequality (5) follows from (40) immediately. It remains to apply Theorem 2.1. □

Proof of Theorem 1.4. Define $m_0, m_1, m_2, \dots$ and $n_1, n_2, \dots$ by

$$m_0 = 0, \quad m_s = 2^{2^s}, \qquad n_s = m_s - m_{s-1}, \qquad s = 1, 2, \dots. \tag{41}$$

It is easy to see that

$$\log n_s \leqslant \log m_s = 2^s\,\log 2, \qquad s = 1, 2, \dots. \tag{42}$$

By Assertion A (see (1.5)), for any $s = 1, 2, \dots$ one can construct on a probability space a sequence of i.i.d. $X_1^{(s)}, \dots, X_{n_s}^{(s)}$ and a sequence of i.i.d. Gaussian $Y_1^{(s)}, \dots, Y_{n_s}^{(s)}$ so that $\mathcal{L}(X_k^{(s)}) = \mathcal{L}(\xi)$, $\mathbf{E}\,Y_k^{(s)} = 0$, $\operatorname{cov} Y_k^{(s)} = \mathbb{I}_d$, and

$$\mathbf{P}\Bigl\{\,c_2\,\Delta_s \geqslant \tau\,d^{3/2}\bigl(c_3\,\log^+ d\,\log n_s + x\bigr)\Bigr\} \leqslant e^{-x}, \qquad x \geqslant 0, \tag{43}$$

where

$$\Delta_s = \max_{1\leqslant r\leqslant n_s}\Bigl|\sum_{k=1}^{r} X_k^{(s)} - \sum_{k=1}^{r} Y_k^{(s)}\Bigr|. \tag{44}$$

It is clear that we can define all the vectors mentioned above on the same probability space so that the collections $\Xi_s = \bigl\{X_1^{(s)}, \dots, X_{n_s}^{(s)};\ Y_1^{(s)}, \dots, Y_{n_s}^{(s)}\bigr\}$, $s = 1, 2, \dots$ are jointly independent. Then we define $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ by

$$\begin{aligned} X_{m_{s-1}+k} &= X_k^{(s)},\\ Y_{m_{s-1}+k} &= Y_k^{(s)},\end{aligned} \qquad k = 1, \dots, n_s, \quad s = 1, 2, \dots. \tag{45}$$

In order to show that these sequences satisfy the assertion of Theorem 1.4, it remains to verify the equality (1.13).

Denoting $L_h^{(s)} = \overline{L}^{(s)}(h)$, $h \in \mathbb{R}^d$, $s = 1, \dots, 2^N$, and using (1.11), (1.12), (2.3) and (2.8), it is easy to see that

$$\widehat{R}_h^{(s)}(t) = \big(\widehat{H}_{h/\sqrt{m}}(t/\sqrt{m}) \big)^m \, \widehat{L}_h^{(s)}(t). \tag{29}$$

The relations (1.10), (24), (28) and (29) imply that

$$\sup_{\|h\|\tau \leqslant 1} \big| \widehat{R}_h^{(s)}(t) \big| \leqslant \Big(1 + \frac{\|t\|}{b_7 \sqrt{m}} \Big)^{-m} \qquad \text{for} \quad \|t\| \geqslant b_7 \sqrt{m} \tag{30}$$

and

$$\sup_{\|h\|\tau \leqslant 1} \sup_{\|t\| \geqslant b_8 \sqrt{m}} \big| \widehat{R}_h^{(s)}(t) \big| \leqslant e^{-m b_{10}}. \tag{31}$$

Using (12), (13), (20) and (30), we get, for $r = 2^{n-1} \cdot 2k + 1, \dots, 2^{n-1}(2k+1)$, $\|t\| \geqslant b_7 \sqrt{2m}$, $t \in \mathbb{R}^{2d}$,

$$\sup_{\|h\|\tau \leqslant 1} \big| \widehat{M}_h^{(r)}(t) \big| \leqslant \min_{\mu = 1,2} \Big(1 + \frac{\|t^{(\mu)}\|}{b_7 \sqrt{m}} \Big)^{-m} \leqslant \Big(1 + \frac{\|t\|}{b_7 \sqrt{2\,m}} \Big)^{-m}. \tag{32}$$

Moreover,

$$\sup_{\|h\|\tau \leqslant 1} \sup_{\|t\| \geqslant b_8 \sqrt{2\,m}} \big| \widehat{M}_h^{(r)}(t) \big| \leqslant e^{-m b_{10}}. \tag{33}$$

Using (2.40), (10), (32) and (33), we see that, for the same r and for $t \in \mathbb{R}^{2d}$, $\|t\| \geqslant b_7 \sqrt{m}$,

$$\sup_{\|h\|\tau\sqrt{2} \leqslant 1} \big| \widehat{Q}_h^{(r)}(t) \big| \leqslant \Big(1 + \frac{\|t\|}{b_7 \sqrt{m}} \Big)^{-m} \tag{34}$$

and

$$\sup_{\|h\|\tau\sqrt{2} \leqslant 1} \sup_{\|t\| \geqslant b_8 \sqrt{m}} \big| \widehat{Q}_h^{(r)}(t) \big| \leqslant e^{-m b_{10}}. \tag{35}$$

It is easy to see that the relations (11), (34) and (35) imply that, for $h \in \mathbb{R}^j$, $\|h\| \sqrt{2}\,\tau < 1$, and for $t \in \mathbb{R}^j$, $\|t\| \geqslant b_7 \sqrt{m}$,

$$\big| \widehat{F}_h(t) \big| \leqslant \Big(1 + \frac{\|t\|}{b_7 \sqrt{m}} \Big)^{-m \cdot 2^{n-1}} \tag{36}$$

and

$$\sup_{\|t\| \geqslant b_8 \sqrt{m}} \big| \widehat{F}_h(t) \big| \leqslant e^{-m b_{10} \cdot 2^{n-1}}. \tag{37}$$

It suffices to prove (36) and (37) for $j = 2\,d$ (for $1 \leqslant j < 2\,d$ one should apply (36) and (37) for $j = 2\,d$ and for $h \in \mathbb{R}^{2d}$, $\|h\| \sqrt{2}\,\tau < 1$, $t \in \mathbb{R}^{2d}$ with $h_m = t_m = 0$, $m = j+1, \dots, 2\,d$).

Note now that the set T defined in (6) satisfies the relation

$$T \subset \Big\{ t \in \mathbb{R}^j : \|t\| \geqslant b_8 \sqrt{m} \Big\} \tag{38}$$

Putting $H \stackrel{\text{def}}{=} \mathcal{L}(\zeta)$, $\psi_h(x) = e^{\langle h,x\rangle}\, p(x)$, $h, x \in \mathbb{R}^d$, and integrating by parts, we see that (for $t \in \mathbb{R}^d$, $t \neq 0$)

$$\begin{aligned}\widehat{H}_h(t) &= \left(\mathbf{E}\, e^{\langle h,\zeta\rangle}\right)^{-1} \int\limits_{\|x\|\leqslant b_1} e^{i\langle t,x\rangle}\, \psi_h(x)\, dx \\ &= -\left(\mathbf{E}\, e^{\langle h,\zeta\rangle}\right)^{-1} \int\limits_{\|x\|\leqslant b_1} \frac{e^{i\langle t,x\rangle}}{i\,\|t\|^2}\, d_t\, \psi_h(x)\, dx, \end{aligned} \tag{21}$$

where $H_h = \overline{H}(h)$. Besides, using (1.9), we see that

$$\sup_{\|x\|\leqslant b_1} \sup_{\|h\|_{b_2}\leqslant 1} \left| d_t\, \psi_h(x)\right| \leqslant b_5\, \|t\|\,. \tag{22}$$

As in the formulation of Theorem 1.3 we denote by b_m different positive quantities depending on H. Note that the quantities depending on the dimension d can be considered as depending on H only as well. From (21) and (22) it follows that

$$\sup_{\|h\|_{b_2}\leqslant 1} \left|\widehat{H}_h(t)\right| \leqslant b_6\, \|t\|^{-1} \tag{23}$$

(note that, by the Jensen inequality, $\mathbf{E}\, e^{\langle h,\zeta\rangle} \geqslant e^{\,\mathbf{E}\,\langle h,\zeta\rangle} = 1$). The inequality (23) implies that

$$\sup_{\|h\|_{b_2}\leqslant 1} \left|\widehat{H}_h(t)\right| \leqslant \left(1 + \frac{\|t\|}{b_7}\right)^{-1} \qquad \text{for} \quad \|t\| \geqslant b_7 = 2\, b_6 \tag{24}$$

and

$$\sup_{\|h\|_{b_2}\leqslant 1} \sup_{\|t\|\geqslant b_7} \left|\widehat{H}_h(t)\right| \leqslant {}^1\!/_2. \tag{25}$$

Since the distributions H_h are absolutely continuous, the relation $\left|\widehat{H}_h(t)\right| = 1$ can be valid for $t = 0$ only. Furthermore, the function $\left|\widehat{H}_h(t)\right|$ considered as a function of two variables h and t is continuous for all $h, t \in \mathbb{R}^d$. Therefore,

$$\sup_{\|h\|_{b_2}\leqslant 1} \sup_{b_8\leqslant\|t\|\leqslant b_7} \left|\widehat{H}_h(t)\right| \leqslant b_9 < 1, \tag{26}$$

where

$$b_8 = \left(4\sqrt{2}\, b_2\, d\right)^{-1} \quad \text{and} \quad b_9 \geqslant {}^1\!/_2. \tag{27}$$

The inequalities (25) and (26) imply that

$$\sup_{\|h\|_{b_2}\leqslant 1} \sup_{\|t\|\geqslant b_8} \left|\widehat{H}_h(t)\right| \leqslant b_9 \stackrel{\text{def}}{=} e^{-b_{10}} < 1. \tag{28}$$

Using (12)–(14), we get

$$\bigl|\widehat{M}_h^{(s)}(t)\bigr| \leqslant \prod_{\mu=1}^{2} \exp\bigl(-b^2 \,\|\, t^{(\mu)} \,\|^2/2\bigr) = \exp\bigl(-b^2 \,\|\, t\,\|^2/2\bigr), \qquad t, h \in \mathbb{R}^{2d}, \quad \|h\|\,\tau < 1. \tag{15}$$

Applying (2.40), (10) and (15) with $t = \mathbb{A}^* u$ and $h = \mathbb{A}^* \gamma$, we see that

$$\bigl|\widehat{Q}_\gamma^{(s)}(u)\bigr| \leqslant \exp\bigl(-b^2 \bigl\|\, \mathbb{A}^* u \,\bigr\|^2/2\bigr) \leqslant \exp\bigl(-b^2 \,\|u\|^2\bigr), \qquad u, \gamma \in \mathbb{R}^{2d}, \quad \|\gamma\|\,\sqrt{2}\,\tau < 1. \tag{16}$$

The relations (11) and (16) imply that

$$\bigl|\widehat{F}_h(t)\bigr| \leqslant \exp\bigl(-b^2 \,\|t\|^2 \cdot 2^{n-2}\bigr), \qquad t, h \in \mathbb{R}^j, \quad \|h\|\,\sqrt{2}\,\tau < 1. \tag{17}$$

It is clear that it suffices to verify (17) for $j = 2d$ (for $1 \leqslant j < 2d$ one should apply (17) for $j = 2d$ and for $t, h \in \mathbb{R}^{2d}$, with $h_m = t_m = 0$, $m = j+1, \ldots, 2d$).

Using (6), (7) and (17), we see that

$$\begin{aligned} \int\limits_T \bigl|\widehat{F}_h(t)\bigr|\,dt &\leqslant \exp\Bigl(-\frac{b^2 \cdot 2^{n-3}}{32\,\tau^2\,j^2}\Bigr) \int\limits_{\mathbb{R}^j} \exp\bigl(-b^2\,\|t\|^2 \cdot 2^{n-3}\bigr)\,dt \\ &= \frac{(2\,\pi)^{j/2}}{(b^2 \cdot 2^{n-2})^{j/2}}\,\exp\Bigl(-\frac{b^2 \cdot 2^n}{2^8\,\tau^2\,j^2}\Bigr) \\ &\leqslant \frac{(2\,\pi)^{j/2}\,\tau^{4j\cdot 2^n}}{(\det \mathbb{D})^{1/2}\,\tau^{2j}} \leqslant \frac{(2\,\pi)^{j/2}\,\tau}{2^{n/2}\,(\det \mathbb{D})^{1/2}}, \end{aligned} \tag{18}$$

if c_1 is small enough. The relations (7) and (18) imply (5). It remains to apply Theorem 2.1 to complete the proof of Theorem 1.1. □

End of the proof of Theorem 1.2. Let now the distributions $\mathcal{L}(\xi_s)$ satisfy the conditions of Theorem 1.2. In this case, according to (2.8) and (2), we have

$$\begin{aligned} \bigl|\widehat{R}_h^{(s)}(t)\bigr| &= \Bigl|\,\frac{\bigl(4 + \tau^2\,(d + \|h\|^2 + 2\,i\,\langle h, t\rangle - \|t\|^2)\bigr)\,\exp\bigl((\|h\|^2 + 2\,i\,\langle h, t\rangle - \|t\|^2)/2\bigr)}{\bigl(4 + \tau^2\,(d + \|h\|^2)\bigr)\,\exp\bigl(\|h\|^2/2\bigr)}\,\Bigr| \\ &\leqslant \bigl(2 + \|t\|^2\bigr)\,\exp\bigl(-\|t\|^2/2\bigr) \\ &\leqslant c_{25}\,\exp\bigl(-\|t\|^2/4\bigr), \qquad\qquad \|h\|\,\tau < 1. \end{aligned} \tag{19}$$

The rest of the proof is omitted. It is similar to that of Theorem 1.1 with $b^2 = 1/2$. The presence of c_{25} in the right-hand side of (19) can be easily compensated by choosing c_1 to be sufficiently small.

End of the proof of Theorem 1.3. Consider the dyadic scheme with

$$\mathcal{L}(\xi_s) = \mathcal{L}(X_s) = L^{(s)}\,P, \qquad s = 1, \ldots, 2^N. \tag{20}$$

where $F = \mathcal{L}\big(\mathbf{U}_{n,k}^{j}\big)$, and σ^2 is the minimal eigenvalue of $\mathbb{D} = \operatorname{cov}\mathbf{U}_{n,k}^{j}$. Note that, according to (3.9), we have

$$\mathbb{D} = 2^n\,\mathbb{I}_j, \quad \sigma^2 = 2^n, \quad \det\mathbb{D} = 2^{nj}. \tag{7}$$

Introduce 2^{n-1} random vectors

$$\mathbf{X}_r = \big(X_r,\, X_{2^{n-1}+r}\big) \in \mathbb{R}^{2d}, \qquad r = 2^{n-1}\cdot 2\,k+1, \ldots, 2^{n-1}\,(2\,k+1). \tag{8}$$

Obviously, these vectors are independent. According to (2.36), (2.37) and (8),

$$\widetilde{\mathbf{U}}_{n,k} = \big(U_{n-1,2k},\, U_{n-1,2k+1}\big) = \sum_{r=2^{n-1}\cdot 2k+1}^{2^{n-1}(2k+1)} \mathbf{X}_r. \tag{9}$$

Denote $R_h^{(s)} = \overline{\mathcal{L}(X_s)}(h)$, for $s = 1, \ldots, 2^N$, $h \in \mathbb{R}^d$, and $M_h^{(r)} \stackrel{\text{def}}{=} \overline{\mathcal{L}(\mathbf{X}_r)}(h)$, $Q_h^{(r)} \stackrel{\text{def}}{=} \overline{\mathcal{L}(\mathbb{A}\mathbf{X}_r)}(h)$, for $r = 2^{n-1}\cdot 2\,k+1, \ldots, 2^{n-1}\,(2\,k+1)$, $h \in \mathbb{R}^{2d}$. As usually, we consider only such h for which these distributions exist. Using (2.8), we see that, for all $t \in \mathbb{R}^{2d}$,

$$\widehat{Q}_h^{(r)}(t) = \frac{\mathbf{E}\,\exp\big(\langle h + it,\, \mathbb{A}\mathbf{X}_r\,\rangle\big)}{\mathbf{E}\,\exp\big(\langle h, \mathbb{A}\mathbf{X}_r\,\rangle\big)} = \frac{\mathbf{E}\,\exp\big(\langle\, \mathbb{A}^*\,h + i\,\mathbb{A}^*\,t, \mathbf{X}_r\,\rangle\big)}{\mathbf{E}\,\exp\big(\langle\, \mathbb{A}^*\,h, \mathbf{X}_r\,\rangle\big)} = \widehat{M}_{\mathbb{A}^*h}^{(r)}(\mathbb{A}^*t). \tag{10}$$

By (2.3) and (9), we have (for $j = 2\,d$)

$$\big|\widehat{F}_h(t)\big| = \prod_{r=2^{n-1}\cdot 2k+1}^{2^{n-1}(2k+1)} \big|\widehat{Q}_h^{(r)}(t)\big|. \tag{11}$$

Split $t = \big(t_1, \ldots, t_{2d}\big) \in \mathbb{R}^{2d}$ as $t = \big(t^{(1)},\, t^{(2)}\big)$, where $t^{(1)} = \big(t_1, \ldots, t_d\big)$ and $t^{(2)} = \big(t_{d+1}, \ldots, t_{2d}\big) \in \mathbb{R}^d$. Using (8), (2.8) and introducing a similar notation for $h \in \mathbb{R}^{2d}$, it is easy to check that

$$\widehat{M}_h^{(r)}(t) = \widehat{R}_{h^{(1)}}^{(r)}\big(t^{(1)}\big)\,\widehat{R}_{h^{(2)}}^{(2^{n-1}+r)}\big(t^{(2)}\big). \tag{12}$$

Note that

$$\|\,t\,\|^2 = \|\,t^{(1)}\,\|^2 + \|\,t^{(2)}\,\|^2\,. \tag{13}$$

End of the proof of Theorem 1.1. Let now the distributions $\mathcal{L}(\xi_s)$ satisfy the conditions of Theorem 1.1. In this case, according to (2.3), we have $R_h^{(s)} = \overline{H}_s(h)\,\overline{G}(h)$. It is well-known that the conjugate distributions $\overline{G}(h)$ of the Gaussian distribution G are also Gaussian with covariance operator $\operatorname{cov}\overline{G}(h) = \operatorname{cov} G = b^2\,\mathbb{I}_d$. Therefore,

$$\big|\widehat{R}_h^{(s)}(t)\big| \leqslant \exp\big(-b^2\,\|t\|^2/2\big), \qquad t, h \in \mathbb{R}^d, \quad \|h\|\,\tau < 1. \tag{14}$$

4. Proofs of Theorems 1.1–1.4

We start the proofs of Theorems 1.1–1.3 with the following common part.

Beginning of the proofs of Theorems 1.1, 1.2 *and* 1.3. At first we shall verify that under the conditions of Theorems 1.2 or 1.3 we have $\mathcal{L}(\xi_k) \in \mathcal{A}_d(\tau)$. For Theorem 1.3 this relation is an immediate consequence of Lemma 3.1, of the completeness of classes $\mathcal{A}_d(\tau)$ with respect to convolution and of the conditions (1.8) and (1.10)–(1.12). In the case of Theorem 1.2 we denote $K = \mathcal{L}(\eta)$. One can easily verify that $\mathbb{B} = \operatorname{cov} K = \gamma^2 \mathbb{I}_d$, where γ^2 is defined by (1.7) and, hence,

$$1 \leqslant \gamma^2 \leqslant 3. \tag{1}$$

Moreover,

$$\varphi(K, z) = \log \mathbf{E}\, e^{\langle z, \eta\rangle} = \log \frac{\bigl(4 + \tau^2 (d + \langle z, \overline{z}\rangle)\bigr)\, \exp\bigl(\langle z, \overline{z}\rangle/2\bigr)}{(4 + \tau^2 d)}, \qquad z \in \mathbb{C}^d. \tag{2}$$

Using (1) and (2), we obtain

$$\bigl|\, d_u d_v^2 \varphi(K, z)\,\bigr| = \Bigl|\, d_u d_v^2 \log\bigl(4 + \tau^2 (d + \langle z, \overline{z}\rangle)\bigr)\,\Bigr| \leqslant c\tau^3 \,\|u\|\, \|v\|^2 \leqslant \|u\|\, \tau \bigl\langle \mathbb{B} v, v\bigr\rangle, \tag{3}$$

for $\|z\|\, \tau \leqslant 1$, if c_1 involved in Assertion A is sufficiently small. This means that $K = \mathcal{L}(\eta) \in \mathcal{A}_d(\tau)$. The relation $\mathcal{L}(\xi_k) = \mathcal{L}\bigl(\eta/\gamma\bigr) \in \mathcal{A}_d(\tau)$, $k = 1, \ldots, n$, follows from (1) and from Lemma 3.1.

The text below is related to Theorems 1.1, 1.2 and 1.3 simultaneously. Without loss of generality we assume that the amount of summands is equal to 2^N with some positive integer N. It suffices to show that the dyadic scheme related to the vectors $\xi_1, \ldots, \xi_{2^N}$ satisfies the conditions of Theorem 2.1 with $\tau^* = \sqrt{2}\,\tau$ instead of τ. According to Lemma 2.1, we can verify the conditions (2.18) and (2.19) for the vectors $\mathbf{U}_{n,k}^j$ and $\mathbf{U}_{N,0}^j$ instead of $\mathbf{U}_{n,k}^{*j}$ and $\mathbf{U}_{N,0}^{*j}$. To this end we shall show that

$$\mathcal{L}\bigl(\mathbf{U}_{n,k}^j\bigr) \in \overline{\mathcal{A}}_j(\sqrt{2}\,\tau, 4) \qquad \text{for} \quad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N, \quad 1 \leqslant j \leqslant 2d. \tag{4}$$

Recall that $\mathbf{U}_{n,k} = \mathbb{A}\widetilde{\mathbf{U}}_{n,k}$, where $\mathbb{A}$ is the linear operator defined by (2.16) and satisfying (2.40). Moreover, $\widetilde{\mathbf{U}}_{n,k} = \bigl(U_{n-1,2k},\, U_{n-1,2k+1}\bigr) \in \mathbb{R}^{2d}$, where $U_{n-1,2k}$ and $U_{n-1,2k+1}$ are independent elements of $\mathbb{R}^d$. The relation $\mathcal{L}(\mathbf{U}_{n,k}) \in \mathcal{A}_{2d}(\sqrt{2}\,\tau)$ can be therefore easily derived from the conditions of Theorems 1.1, 1.2 and 1.3 with the help of Lemmas 2.1, 3.1 and 3.2 (see (2.40)) if we take into account the completeness of classes $\mathcal{A}_d(\tau)$ with respect to convolution and their monotonicity with respect to τ. It is easy to see that $\mathbf{U}_{n,k}^j = \overline{\mathbb{P}}_j\, \mathbf{U}_{n,k}$, where the projector $\overline{\mathbb{P}}_j : \mathbb{R}^{2d} \to \mathbb{R}^j$ can be considered as a linear operator with $\|\overline{\mathbb{P}}_j\| = 1$ (see (2.34)). Applying Lemma 3.1 again, we obtain the relations $\mathcal{L}(\mathbf{U}_{n,k}^j) \in \mathcal{A}_j(\sqrt{2}\,\tau)$, $1 \leqslant j \leqslant 2d$.

It remains to verify that, for $h \in \mathbb{R}^j$, $\|h\|\, \sqrt{2}\,\tau < 1$, the following inequality hold:

$$\int\limits_T \bigl|\, \widehat{F}_h(t)\bigr|\, dt \leqslant \frac{(2\,\pi)^{j/2}\, \sqrt{2}\,\tau\, j^{3/2}}{\sigma\, (\det \mathbb{D})^{1/2}}, \tag{5}$$

$$T = \Bigl\{ t \in \mathbb{R}^j : 4\, \|t\|\, \sqrt{2}\,\tau\, j \geqslant 1 \Bigr\}, \tag{6}$$

Putting now $t^* = (8\,c_{23}\,d^{3/2}\,\tau)^{-1}$ and $t = t^* \cdot c_{23}\,d^{3/2}\,\tau = 1/8$, taking into account that the random variables $W^{(L)}$, $W_{(L)}$, $\dots$, $W_{(N-1)}$ are independent and applying (60), (64) and (65), we obtain

$$\begin{aligned}
&\mathbf{P}\left\{\left\{\omega : \left| S_j - T_j\right| \geqslant x/2\right\} \cap A\right\} \\
&\quad\leqslant \mathbf{P}\left\{ c_{23}\,d^{3/2}\,\tau\left(N+1+W^{(L)}+\sum_{n=L}^{N-1} W_{(n)}\right) \geqslant x/2 \right\} \\
&\quad\leqslant \mathbf{P}\left\{ t\left(W^{(L)}+\sum_{n=L}^{N-1} W_{(n)}\right) \geqslant t^* x/2 - t\,(N+1)\right\} \\
&\quad\leqslant \mathbf{E}\,\exp\left(t\left(W^{(L)}+\sum_{n=L}^{N-1} W_{(n)}\right)\right) \Big/ \exp\left(t^* x/2 - t\,(N+1)\right) \\
&\quad= \mathbf{E}\,\exp\left(t W^{(L)}\right) \prod_{n=L}^{N-1} \mathbf{E}\,\exp\left(t W_{(n)}\right) \Big/ \exp\left(t^* x/2 - t\,(N+1)\right) \\
&\quad\leqslant (3\,d)^{N+1}\,\exp\left(\frac{N+1}{8} - \frac{x}{16\,c_{23}\,d^{3/2}\,\tau}\right). \qquad (66)
\end{aligned}$$

From (35), (45) and (66) it follows that

$$\mathbf{P}\left\{\Delta_5 \geqslant x/2,\ A\right\} \leqslant 2^N \cdot (3\,d)^{N+1}\,\exp\left(\frac{N+1}{8} - \frac{x}{16\,c_{23}\,d^{3/2}\,\tau}\right). \qquad (67)$$

Using (31), (39), (42)–(44) and (67), we obtain that

$$\mathbf{P}\left\{\Delta \geqslant x\right\} \leqslant (19\,d)^{N+1}\,\exp\left(-\frac{x}{c_{24}\,d^{3/2}\,\tau}\right), \qquad x \geqslant 0, \qquad (68)$$

where we can take $c_{24} = \max\left\{16\,c_{23},\ c_{19}^{-1},\ c_{20}^{-1},\ c_{21}^{-1},\ 2\right\}$. Let the quantities ε, $x_0 > 0$ be defined by the relations

$$\varepsilon = \frac{1}{2\,c_{24}\,d^{3/2}\,\tau} \leqslant \frac{1}{4\,\tau}, \qquad e^{\varepsilon x_0} = (19\,d)^{N+1}. \qquad (69)$$

Integrating by parts and using (68) and (69), we obtain

$$\mathbf{E}\,e^{\varepsilon\Delta} = \int_0^\infty \varepsilon\,e^{\varepsilon x}\ \mathbf{P}\left\{\Delta \geqslant x\right\} dx + 1,$$

$$\int_0^{x_0} \varepsilon\,e^{\varepsilon x}\ \mathbf{P}\left\{\Delta \geqslant x\right\} dx \leqslant \int_0^{x_0} \varepsilon\,e^{\varepsilon x}\,dx = e^{\varepsilon x_0} - 1 = (19\,d)^{N+1} - 1,$$

$$\int_{x_0}^\infty \varepsilon\,e^{\varepsilon x}\ \mathbf{P}\left\{\Delta \geqslant x\right\} dx \leqslant \int_{x_0}^\infty \varepsilon\,e^{-\varepsilon(x-x_0)}\,dx = 1,$$

and, hence,

$$\mathbf{E}\,e^{\varepsilon\Delta} \leqslant (19\,d)^{N+1} + 1 \leqslant (20\,d)^{N+1}.$$

Together with (24) and (69), this completes the proof of Theorem 2.1. $\square$

From (49) and (57) it follows that for $\omega \in A$ we have

$$\left| S_j - T_j \right| \leqslant c_{23}\, d^{3/2}\, \tau \left(N + 1 + \frac{\left| U^{(L)} \right|^2}{2^L} + \sum_{n=L}^{N-1} \frac{\left| U_{(n)} \right|^2}{2^n} \right). \tag{58}$$

Denote (for $0 \leqslant n \leqslant N,\ \ 0 \leqslant l < 2^{N-n}$)

$$W_{n,l} = \begin{cases} 2^{-n} \left| U_{n,l} \right|^2, & \text{if} \quad \left| U_{n,l} \right| \leqslant y \cdot 2^n, \\ 0, & \text{otherwise.} \end{cases} \tag{59}$$

Let us show that

$$\mathbf{E} \exp\left(t\, W_{n,l}\right) \leqslant 2\, d + 1, \qquad \text{for} \quad 0 \leqslant t \leqslant \frac{1}{8}. \tag{60}$$

Indeed, integrating by parts, we obtain

$$\begin{aligned} \mathbf{E} \exp(t W_{n,l}) &= 1 + \int\limits_0^{y^2 \cdot 2^n} t \exp(t\, u)\ \mathbf{P}\left\{ W_{n,l} \geqslant u \right\} du \\ &\leqslant 1 + \frac{1}{8} \int\limits_0^{y^2 \cdot 2^n} \exp\left(u/8\right)\ \mathbf{P}\left\{ \left| U_{n,l} \right| \geqslant 2^{n/2} \sqrt{u} \right\} du. \end{aligned} \tag{61}$$

Taking into account (2.37), (25) and using Lemmas 3.1 and 3.3, we obtain that

$$\begin{aligned} \mathbf{P}\left\{ \left| U_{n,l} \right| \geqslant 2^{n/2} \sqrt{u} \right\} &\leqslant 2\, d\, \exp\left(- \min\left\{ \frac{2^n\, u}{4 \cdot 2^n},\ \frac{2^{n/2} \sqrt{u}}{4\, \tau} \right\} \right) \\ &\leqslant 2\, d\, \exp\left(- \min\left\{ \frac{u}{4},\ \frac{u}{4\, y\, \tau} \right\} \right) \\ &= 2\, d\, \exp\left(- \frac{u}{4} \right), \end{aligned} \tag{62}$$

if $0 \leqslant u \leqslant y^2 \cdot 2^n$. The relation (60) immediately follows from (61) and (62).

The relations (47), (48) and (59) imply that, for $L \leqslant n \leqslant N,\ \ 0 \leqslant l < 2^{N-n}$, $\omega \in A$,

$$2^{-n} \left| U_{n,l} \right|^2 = W_{n,l}. \tag{63}$$

Thus, according to (50), we can rewrite (58) in the form

$$\left| S_j - T_j \right| \leqslant c_{23}\, d^{3/2}\, \tau \left(N + 1 + W^{(L)} + \sum_{n=L}^{N-1} W_{(n)} \right), \qquad \omega \in A, \tag{64}$$

where

$$W^{(L)} = W_{L, l_{L,j}}, \qquad W_{(n)} = W_{n, \widetilde{l}_{n,j}}, \tag{65}$$

(it is easy to see that $l_{n-1,j}$ can be equal either to $2l_{n,j}$ or to $2l_{n,j}+1$, for given $l_{n,j}$). In other words, $U^{(n)}$, $L \leqslant n \leqslant N$, is the sum over the block of 2^n summands which contains X_j. The sum $U_{(n)}$ does not contain X_j and

$$U^{(n+1)} = U^{(n)} + U_{(n)}, \qquad L \leqslant n < N \tag{52}$$

(see (2.37)). The equality (52) implies

$$U^{(n)} = U^{(L)} + \sum_{s=0}^{n-L-1} U_{(L+s)}, \qquad L \leqslant n \leqslant N. \tag{53}$$

It is important that all summands in the right-hand side of (53) are the sums of disjoint blocks of independent summands. Therefore, they are independent.

Put $\beta = 1/\sqrt{2}$. Then, using (53) and the Hölder inequality, one can easily derive that, for $L \leqslant n \leqslant N$,

$$\left|U^{(n)}\right|^2 \leqslant c_{22}\left(\beta^{-(n-L)}\left|U^{(L)}\right|^2 + \sum_{s=0}^{n-L-1}\beta^{-(n-L-1)+s}\left|U_{(L+s)}\right|^2\right), \tag{54}$$

with $c_{22} = \sum\limits_{j=0}^{\infty}\beta^j = \dfrac{\sqrt{2}}{\sqrt{2}-1}$. It is easy to see that

$$\sum_{n=L}^{N} 2^{-n}\,\beta^{-(n-L)}\left|U^{(L)}\right|^2 \leqslant c_{22}\cdot 2^{-L}\left|U^{(L)}\right|^2. \tag{55}$$

Moreover,

$$\begin{aligned}\sum_{n=L+1}^{N}\sum_{s=0}^{n-L-1} & 2^{-n}\beta^{-(n-L-1)+s}\left|U_{(L+s)}\right|^2 \\ &= \sum_{s=0}^{N-L-1}\sum_{n=L+1+s}^{N} 2^{-n}\beta^{-(n-L-1)+s}\left|U_{(L+s)}\right|^2 \\ &\leqslant c_{22}\sum_{s=0}^{N-L-1} 2^{-(L+1+s)}\left|U_{(L+s)}\right|^2.\end{aligned} \tag{56}$$

It is clear that the inequalities (54)–(56) imply

$$\begin{aligned}2^{-N}\left|U_{N,0}\right|^2 & + \sum_{n=L}^{N-1} 2^{-n}\left(\left|U^{(n)}\right|^2 + \left|U_{(n)}\right|^2\right) \\ &\leqslant c_{22}\left(\frac{\left|U^{(L)}\right|^2}{2^L} + \sum_{s=0}^{N-L-1}\frac{\left|U_{(L+s)}\right|^2}{2^{L+1+s}}\right) + \sum_{n=L}^{N-1}\frac{\left|U_{(n)}\right|^2}{2^n} \\ &\leqslant c\left(\frac{\left|U^{(L)}\right|^2}{2^L} + \sum_{n=L}^{N-1}\frac{\left|U_{(n)}\right|^2}{2^n}\right).\end{aligned} \tag{57}$$

It remains to estimate $\mathbf{P}\left\{\Delta_5 \geqslant x/2,\ A\right\}$ in both cases: $L = M$ and $L = 0$ (see (39) and (42)–(44)). Let L defined by (32) be arbitrary. Fix an integer k satisfying $1 \leqslant k \leqslant 2^{N-L}$ and denote for simplicity

$$j = j(k) \stackrel{\text{def}}{=} k \cdot 2^L. \tag{45}$$

By Corollary 2.1, we have

$$\left| S_{k\cdot 2^L} - T_{k\cdot 2^L} \right| = \left| S_j - T_j \right| \leqslant \left| U_{N,0} - V_{N,0} \right| + \frac{1}{2} \sum_{n=L+1}^{N} \left| \widetilde{U}_{n,l_{n,j}} - \widetilde{V}_{n,l_{n,j}} \right|, \tag{46}$$

where $l_{n,j}$ are integers, defined by $l_{n,j} \cdot 2^n < j \leqslant \left(l_{n,j}+1\right) \cdot 2^n$ (see (2.42)).

By virtue of (25) and (36), for $\omega \in A$ we have

$$\left| U_{L,l} \right| < y \cdot 2^L = \frac{c_{18} \cdot 2^L}{d^{3/2}\,\tau} \leqslant \frac{\min\{c_{14},\, c_{17}\} \cdot 2^L}{d^{3/2}\,\tau}, \qquad 0 \leqslant l < 2^{N-L}, \tag{47}$$

and, by (2.35)–(2.37), $U_{L,l}$ are sums over blocks consisting of 2^L summands. Moreover, $U_{n,l}$ (resp. $\widetilde{U}_{n,l}$), $L+1 \leqslant n \leqslant N$, $0 \leqslant l < 2^{N-n}$, are sums (resp. differences) of two sums over blocks containing each 2^{n-1} summands. These sums and differences can be represented as linear combinations (with coefficients ± 1) of 2^{n-L} sums over blocks containing each 2^L summands and satisfying (47). Therefore, for $\omega \in A$, $L+1 \leqslant n \leqslant N$, $0 \leqslant l < 2^{N-n}$ we have (see (2.32) and (2.34))

$$\left| \mathbf{U}_{n,l} \right| = \max\left\{ \left| U_{n,l} \right|,\ \left| \widetilde{U}_{n,l} \right| \right\} \leqslant 2^{n-L}\, y \cdot 2^L = y \cdot 2^n \leqslant \frac{\min\{c_{14},\, c_{17}\} \cdot 2^n}{d^{3/2}\,\tau}. \tag{48}$$

Using (48), we see that if $\omega \in A$, the conditions of Lemma 3.6 are satisfied for τ, $U_{N,0}$ and $\mathbf{U}_{n,l}$, if $L+1 \leqslant n \leqslant N$, $0 \leqslant l < 2^{N-n}$. By (46), (48) and by Lemma 3.6, for $\omega \in A$ we have

$$\begin{aligned} \left| S_j - T_j \right| &\leqslant c_{13}\, d^{3/2}\,\tau \left(1 + 2^{-N} \left| U_{N,0} \right|^2\right) \\ &\quad + \sum_{n=L+1}^{N} c_{16}\, d^{3/2}\,\tau \left(1 + 2^{-n} \max\left\{ \left| U_{n,l_{n,j}} \right|^2,\ \left| \widetilde{U}_{n,l_{n,j}} \right|^2 \right\}\right) \\ &\leqslant c d^{3/2}\,\tau \left(N + 1 + 2^{-N} \left| U_{N,0} \right|^2 + \sum_{n=L}^{N-1} 2^{-n} \left(\left| U^{(n)} \right|^2 + \left| U_{(n)} \right|^2 \right) \right), \end{aligned} \tag{49}$$

where

$$U^{(n)} = U_{n,l_{n,j}}, \qquad U_{(n)} = U_{n,\widetilde{l}_{n,j}}, \tag{50}$$

and

$$\widetilde{l}_{n-1,j} = \begin{cases} 2\,l_{n,j}, & \text{if} \quad l_{n-1,j} = 2\,l_{n,j} + 1, \\ 2\,l_{n,j} + 1, & \text{if} \quad l_{n-1,j} = 2\,l_{n,j}, \end{cases} \qquad L < n \leqslant N \tag{51}$$

Introduce the event

$$A = \Big\{ \omega : \ \big| U_{L,k} \big| < y \cdot 2^L, \ 0 \leqslant k < 2^{N-L} \Big\} \tag{36}$$

(we assume that all considered random vectors are measurable mappings of $\omega \in \Omega$). For the complementary event we use the notation $\overline{A} = \Omega \setminus A$.

We consider separately two possible cases: $L = M$ and $L = 0$. Let $L = M$. It is evident that in this case

$$\Delta \leqslant \Delta_3 + \Delta_4 + \Delta_5. \tag{37}$$

Moreover, by virtue of (2.37), (26), (33) and (36), we have

$$\overline{A} \subset \Big\{ \omega : \Delta_3 \geqslant x/4 \Big\}. \tag{38}$$

From (37) and (38) it follows that

$$\mathbf{P}\Big\{ \Delta \geqslant x \Big\} \leqslant \mathbf{P}\Big\{ \Delta_3 \geqslant x/4 \Big\} + \mathbf{P}\Big\{ \Delta_4 \geqslant x/4 \Big\} + \mathbf{P}\Big\{ \Delta_5 \geqslant x/2, \ A \Big\}. \tag{39}$$

Using Lemmas 3.5, 3.1 and 3.3, the completeness of classes $\mathcal{A}_d(\tau)$ with respect to convolution and the relations (25) and (26), we obtain, for $0 \leqslant k < 2^{N-L}$, that $2^L = 2^M \leqslant x/2\,y$ and

$$\begin{aligned} \mathbf{P}\Big\{ \max_{1 \leqslant l \leqslant 2^L} \big| S_{k \cdot 2^L + l} - S_{k \cdot 2^L} \big| \geqslant x/4 \Big\} &\leqslant 3 \max_{1 \leqslant l \leqslant 2^L} \mathbf{P}\Big\{ \big| S_{k \cdot 2^L + l} - S_{k \cdot 2^L} \big| \geqslant x/12 \Big\} \\ &\leqslant 6\,d \, \exp\Big(- \min\Big\{ \frac{x^2}{576 \cdot 2^L}, \, \frac{x}{48\,\tau} \Big\} \Big) \\ &\leqslant 6\,d \, \exp\Big(- \frac{c_{20}\, x}{d^{3/2}\,\tau} \Big). \end{aligned} \tag{40}$$

Since all d-dimensional Gaussian distributions belong to classes $\mathcal{A}_d(\tau)$ for all $\tau \geqslant 0$, we immediately obtain that

$$\mathbf{P}\Big\{ \max_{1 \leqslant l \leqslant 2^L} \big| T_{k \cdot 2^L + l} - T_{k \cdot 2^L} \big| \geqslant x/4 \Big\} \leqslant 6\,d \, \exp\Big(- \frac{c_{20}\, x}{d^{3/2}\,\tau} \Big). \tag{41}$$

From (33), (34), (40) and (41) it follows that

$$\mathbf{P}\Big\{ \Delta_3 \geqslant x/4 \Big\} + \mathbf{P}\Big\{ \Delta_4 \geqslant x/4 \Big\} \leqslant 2^N \cdot 12\,d \, \exp\Big(- \frac{c_{20}\, x}{d^{3/2}\,\tau} \Big). \tag{42}$$

Assume that $L = 0$. Then, according to (24) and (35), $\Delta = \Delta_5$ and, hence, we have the rough bound

$$\mathbf{P}\Big\{ \Delta \geqslant x \Big\} \leqslant \mathbf{P}\big\{ \overline{A} \big\} + \mathbf{P}\Big\{ \Delta_5 \geqslant x/2, \ A \Big\}. \tag{43}$$

In this case $U_{L,k} = X_{k+1}$, $2^L = 1 \geqslant 2^M$, $y > x/4$ (see (25), (26) and (32)). Therefore, by (36) and by Lemmas 3.1 and 3.3,

$$\begin{aligned} \mathbf{P}\big\{ \overline{A} \big\} &\leqslant \sum_{k=0}^{2^N - 1} \mathbf{P}\Big\{ \big| U_{L,k} \big| \geqslant y \cdot 2^L \Big\} = \sum_{k=1}^{2^N} \mathbf{P}\Big\{ \big| X_k \big| \geqslant y \Big\} \\ &\leqslant 2^{N+1}\, d \, \exp\Big(- \min\Big\{ \frac{y^2}{4}, \, \frac{y}{4\,\tau} \Big\} \Big) \leqslant 2^{N+1}\, d \, \exp\Big(- \min\Big\{ \frac{x\,y}{16}, \, \frac{x}{16\,\tau} \Big\} \Big) \\ &\leqslant 2^{N+1}\, d \, \exp\Big(- \frac{c_{21}\, x}{d^{3/2}\,\tau} \Big). \end{aligned} \tag{44}$$

Proof of Theorem 2.1. Let X_k, $k = 1, \dots, 2^N$, denote the vectors constructed by the dyadic procedure described in (2.22)–(2.36). Denote

$$\Delta = \Delta(2^N) = \max_{1 \leqslant k \leqslant 2^N} \big| S_k - T_k \big|, \tag{24}$$

$$c_5 = \min\big\{ c_9,\, c_{15} \big\}, \qquad c_{18} = \min\big\{ c_{14},\, c_{17},\, 1 \big\}, \qquad y \stackrel{\text{def}}{=} \frac{c_{18}}{d^{3/2}\,\tau} \leqslant \frac{1}{\tau}, \tag{25}$$

fix some $x > 0$ and choose the integer M such that

$$x < 4\,y \cdot 2^M \leqslant 2\,x. \tag{26}$$

We shall estimate $\mathbf{P}\big\{ \Delta \geqslant x \big\}$. Consider separately two possible cases: $M \geqslant N$ and $M < N$. Let, at first, $M \geqslant N$. Denote

$$\Delta_1 = \max_{1 \leqslant k \leqslant 2^N} \big| S_k \big|, \qquad \Delta_2 = \max_{1 \leqslant k \leqslant 2^N} \big| T_k \big|. \tag{27}$$

It is easy to see that $\Delta \leqslant \Delta_1 + \Delta_2$ and, hence,

$$\mathbf{P}\big\{ \Delta \geqslant x \big\} \leqslant \mathbf{P}\big\{ \Delta_1 \geqslant x/2 \big\} + \mathbf{P}\big\{ \Delta_2 \geqslant x/2 \big\}. \tag{28}$$

Taking into account the completeness of classes $\mathcal{A}_d(\tau)$ with respect to convolution, applying Lemmas 3.5, 3.1 and 3.3 and using (25) and (26), we obtain that $2^N \leqslant 2^M \leqslant x/2\,y$ and

$$\begin{aligned} \mathbf{P}\big\{ \Delta_1 \geqslant x/2 \big\} \leqslant 3 \max_{1 \leqslant k \leqslant 2^N} \mathbf{P}\big\{ \big| S_k \big| \geqslant x/6 \big\} &\leqslant 6\,d\, \exp\Big(- \min\Big\{ \frac{x^2}{144 \cdot 2^N},\, \frac{x}{24\,\tau} \Big\} \Big) \\ &\leqslant 6\,d\, \exp\Big(- \frac{c_{19}\, x}{d^{3/2}\,\tau} \Big). \end{aligned} \tag{29}$$

Since all d-dimensional Gaussian distributions belong to all classes $\mathcal{A}_d(\tau)$, $\tau \geqslant 0$, we automatically obtain that

$$\mathbf{P}\big\{ \Delta_2 \geqslant x/2 \big\} \leqslant 6\,d\, \exp\Big(- \frac{c_{19}\, x}{d^{3/2}\,\tau} \Big). \tag{30}$$

From (28)–(30) it follows in the case $M \geqslant N$ that

$$\mathbf{P}\big\{ \Delta \geqslant x \big\} \leqslant 12\,d\, \exp\Big(- \frac{c_{19}\, x}{d^{3/2}\,\tau} \Big). \tag{31}$$

Let now $M < N$. Denote

$$L = \max\big\{ 0,\, M \big\} \tag{32}$$

and

$$\Delta_3 = \max_{0 \leqslant k < 2^{N-L}} \max_{1 \leqslant l \leqslant 2^L} \big| S_{k \cdot 2^L + l} - S_{k \cdot 2^L} \big|, \tag{33}$$

$$\Delta_4 = \max_{0 \leqslant k < 2^{N-L}} \max_{1 \leqslant l \leqslant 2^L} \big| T_{k \cdot 2^L + l} - T_{k \cdot 2^L} \big|, \tag{34}$$

$$\Delta_5 = \max_{1 \leqslant k \leqslant 2^{N-L}} \big| S_{k \cdot 2^L} - T_{k \cdot 2^L} \big|. \tag{35}$$

if $\frac{\tau}{2^{N/2}} \leqslant c_9$, $\left|U_{N,0}^{(1)}\right| \leqslant \frac{c_{11}\cdot 2^N}{\tau}$. Furthermore,

$$\left|U_{N,0}^{(j)} - V_{N,0}^{(j)}\right| \leqslant c_{12}\,\tau\left(j^{3/2} + j^{3/2}\,\frac{\left|U_{N,0}^{j-1}\right|}{2^{N/2}}\left(1+\frac{\left|U_{N,0}^{(j)} - y_j\right|}{2^{N/2}}\right) + \frac{\left|U_{N,0}^{(j)} - y_j\right|^2}{2^N}\right) + |y_j|, \tag{14}$$

if

$$\frac{\tau\, j^{3/2}}{2^{N/2}} \leqslant c_9, \quad \frac{\left|U_{N,0}^{j-1}\right|}{2^{N/2}} \leqslant \frac{c_{10}\cdot 2^{N/2}}{j^{3/2}\tau}, \quad \left|U_{N,0}^{(j)} - y_j\right| \leqslant \frac{c_{11}\cdot 2^N}{j\,\tau}, \qquad 2 \leqslant j \leqslant d, \tag{15}$$

where

$$|y_j| \leqslant c_8\,\tau\, j\,\frac{\left|U_{N,0}^{j-1}\right|^2}{2^N}, \qquad 2 \leqslant j \leqslant d. \tag{16}$$

Obviously,

$$\left|U_{N,0}^{(1)}\right| \leqslant \max\left\{\left|U_{N,0}^{j-1}\right|,\ \left|U_{N,0}^{(j)}\right|\right\} = \left|U_{N,0}^{j}\right| \leqslant \left|U_{N,0}\right|, \qquad 2 \leqslant j \leqslant d, \tag{17}$$

see (2.31) and (10). Using (13), (14), (16) and (17), we see that one can choose c_{13} to be so large and c_{14} to be so small that

$$\left|U_{N,0}^{(j)} - V_{N,0}^{(j)}\right| \leqslant c_{13}\,d^{3/2}\,\tau\left(1 + 2^{-N}\left|U_{N,0}\right|^2\right), \tag{18}$$

if $\frac{\tau\, d^{3/2}}{2^{N/2}} \leqslant c_9$, $\left|U_{N,0}\right| \leqslant \frac{c_{14}\cdot 2^N}{d^{3/2}\tau}$, $1 \leqslant j \leqslant d$. The inequality (7) immediately follows from (18), (2.23) and (2.31).

Now we shall prove item b). According to Lemma 2.1, Remark 2.1, (2.18), (2.31) and (9), the distributions $\mathcal{L}(\mathbf{U}_{n,k}^j)$, $j = d+1,\dots,2\,d$, satisfy in the j-dimensional case the conditions of Lemma 3.4 with $\sigma^2 = 2^n$, $\mathbb{B} = \operatorname{cov}\mathbf{U}_{n,k}^{j-1} = 2^n\,\mathbb{I}_{j-1}$.

Using (2.33) and applying Lemmas 3.4 and 3.7, we obtain that

$$\left|\mathbf{U}_{n,k}^{(j)} - \mathbf{V}_{n,k}^{(j)}\right| \leqslant c_{12}\,\tau\left(j^{3/2} + j^{3/2}\,\frac{\left|\mathbf{U}_{n,k}^{j-1}\right|}{2^{n/2}}\left(1+\frac{\left|\mathbf{U}_{n,k}^{(j)} - y_j\right|}{2^{n/2}}\right) + \frac{\left|\mathbf{U}_{n,k}^{(j)} - y_j\right|^2}{2^n}\right) + |y_j|, \tag{19}$$

if

$$\frac{\tau\, j^{3/2}}{2^{n/2}} \leqslant c_9, \quad \frac{\left|\mathbf{U}_{n,k}^{j-1}\right|}{2^{n/2}} \leqslant \frac{c_{10}\cdot 2^{n/2}}{j^{3/2}\tau}, \quad \left|\mathbf{U}_{n,k}^{(j)} - y_j\right| \leqslant \frac{c_{11}\cdot 2^n}{j\,\tau}, \tag{20}$$

where

$$|y_j| \leqslant c_8\,\tau\, j\,\frac{\left|\mathbf{U}_{n,k}^{j-1}\right|^2}{2^n}, \qquad d+1 \leqslant j \leqslant 2\,d. \tag{21}$$

Obviously,

$$\max\left\{\left|\mathbf{U}_{n,k}^{j-1}\right|,\ \left|\mathbf{U}_{n,k}^{(j)}\right|\right\} = \left|\mathbf{U}_{n,k}^{j}\right| \leqslant \left|\mathbf{U}_{n,k}\right|, \tag{22}$$

see (2.34). Using (19), (21) and (22), we see that one can choose c_{15} and c_{17} to be so small and c_{16} to be so large that

$$\left|\mathbf{U}_{n,k}^{(j)} - \mathbf{V}_{n,k}^{(j)}\right| \leqslant c_{16}\,d^{3/2}\,\tau\left(1 + 2^{-n}\left|\mathbf{U}_{n,k}\right|^2\right) \tag{23}$$

if $\frac{\tau\, d^{3/2}}{2^{n/2}} \leqslant c_{15}$, $\left|\mathbf{U}_{n,k}\right| \leqslant \frac{c_{17}\cdot 2^n}{d^{3/2}\tau}$, $d+1 \leqslant j \leqslant 2\,d$. The inequality (8) immediately follows from (23), (2.24), (2.25) and (2.34). □

Lemma 3.7. *Let $F(\cdot)$ denote a continuous distribution function and $G(\cdot)$ an arbitrary distribution function satisfying for $z \in B \in \mathfrak{B}_1$ the inequality*

$$G\big(z - f(z)\big) < F(z + w) < G\big(z + f(z)\big)$$

with some $f : B \to \mathbb{R}^1$ and $w \in \mathbb{R}^1$. Let $\eta \in \mathbb{R}^1$, $0 < G(\eta) < 1$ and $\xi = F^{-1}\big(G(\eta)\big)$, where $F^{-1}(x) = \sup\big\{u : F(u) \leqslant x\big\}$, $0 < x < 1$. Then

$$|\,\xi - \eta\,| < f(\xi - w) + |\,w\,|, \qquad \text{if} \quad \xi - w \in B.$$

Proof. Put $\zeta = \xi - w$. The continuity of F implies that $F\big(F^{-1}(x)\big) \equiv x$, for $0 < x < 1$. Therefore,

$$\zeta \in B \implies G\big(\zeta - f(\zeta)\big) < F(\xi) = G(\eta) \implies \zeta - f(\zeta) < \eta \implies \xi - \eta < f(\zeta) + w$$

and

$$\zeta \in B \implies G(\eta) = F(\xi) < G\big(\zeta + f(\zeta)\big) \implies \eta < \zeta + f(\zeta) \implies \eta - \xi < f(\zeta) - w.$$

This completes the proof of the lemma. □

Proof of Lemma 3.6. At first we note that the conditions of Theorem 2.1 imply that

$$\operatorname{cov} \mathbf{U}_{n,k} = 2^n \mathbb{I}_{2d}, \qquad \text{for} \quad 1 \leqslant n \leqslant N, \quad 0 \leqslant k < 2^{N-n},$$

and, hence (see (2.28)),

$$\operatorname{cov} \mathbf{U}^j_{n,k} = 2^n \mathbb{I}_j, \qquad \text{for} \quad 1 \leqslant j \leqslant 2\,d. \tag{9}$$

Let us prove the assertion a). Introduce the vectors

$$U^j_{N,0} = \big(U^{(1)}_{N,0}, \dots, U^{(j)}_{N,0}\big), \qquad V^j_{N,0} = \big(V^{(1)}_{N,0}, \dots, V^{(j)}_{N,0}\big) \tag{10}$$

consisting of the first j coordinates of the vectors $U_{N,0}$, $V_{N,0}$ respectively. By (9), (2.32) and (2.34),

$$U_{N,0} = \overline{\mathbb{P}}_d\, \mathbf{U}_{N,0} \tag{11}$$

and

$$U^j_{N,0} = \mathbf{U}^j_{N,0}, \qquad \operatorname{cov} U^j_{N,0} = 2^n \mathbb{I}_j, \qquad \text{for} \quad 1 \leqslant j \leqslant d. \tag{12}$$

Moreover, according to Lemma 2.1, Remark 2.1, (12) and (2.19), the distributions $\mathcal{L}(U^j_{N,0})$, $j = 1, \dots, d$, satisfy in the j-dimensional case the conditions of Lemma 3.4 with $\sigma^2 = 2^N$ and $\mathbb{B} = \operatorname{cov} U^{j-1}_{N,0} = 2^N \mathbb{I}_{j-1}$ (the last equality for $j \geqslant 2$).

Taking into account (2.29) and applying Lemmas 3.4 and 3.7, we obtain that

$$\big|\,U^{(1)}_{N,0} - V^{(1)}_{N,0}\,\big| \leqslant c_{12}\,\tau \left(1 + \frac{\big|\,U^{(1)}_{N,0}\,\big|^2}{2^N}\right), \tag{13}$$

for $\dfrac{\tau\, d^{3/2}}{\sigma} \leqslant c_9$, $\left|\mathbb{B}^{-1/2}x\right| \leqslant \dfrac{c_{10}\,\sigma}{d^{3/2}\,\tau}$, $|z| \leqslant \dfrac{c_{11}\,\sigma^2}{d\,\tau}$, *where*

$$\gamma(z) = c_{12}\,\tau\left(d^{3/2} + d\,\delta\left(1 + \frac{|z|}{\sigma}\right) + \frac{z^2}{\sigma^2}\right), \qquad \delta = \left\|\mathbb{B}^{-1/2}x\right\|. \tag{3}$$

b) *The assertion* a) *remains valid for* $d = 1$ *with* $F(z\,|\,x) = \mathbf{P}\left\{\xi_1 < z\right\}$ *and* $y = \delta = 0$ *without any restrictions on* $\mathbb{B}$, $\overline{\mathbb{P}}_{d-1}\,\xi$ *and* x.

Remark 3.1. In Zaitsev (1995, 1996a) the formulation of Lemma 3.4 is in some sense weaker, see Zaitsev (1995, 1996a, Lemmas 6.1 and 8.1). In particular, instead of the conditions

$$\mathcal{L}(\xi) \in \overline{\mathcal{A}}_d(\tau, 4) \qquad \text{and} \qquad \mathcal{L}(\overline{\mathbb{P}}_{d-1}\,\xi) \in \overline{\mathcal{A}}_{d-1}(\tau, 4) \tag{4}$$

the stronger conditions

$$\mathcal{L}(\xi) \in \mathcal{A}_d^*(\tau, 4, 4) \qquad \text{and} \qquad \mathcal{L}(\overline{\mathbb{P}}_{d-1}\,\xi) \in \mathcal{A}_{d-1}^*(\tau, 4, 4) \tag{5}$$

are used. However, in the proof of (1) and (2) only the conditions (4) are applied. The conditions (5) are necessary for the investigation of quantiles of conditional distributions corresponding to random vectors having coinciding moments up to third order which has been done in Zaitsev (1995, 1996a) simultaneously with the proof of (1) and (2).

Lemma 3.5. *Let* $S_k = X_1 + \cdots + X_k$ *be sums of independent random vectors* $X_j \in \mathbb{R}^d$ *and let* $q(\cdot)$ *be a semi-norm in* $\mathbb{R}^d$. *Then*

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n} q(S_k) > 3\,t\right\} \leqslant 3\max_{1\leqslant k\leqslant n}\mathbf{P}\left\{q(S_k) > t\right\}, \qquad t \geqslant 0. \tag{6}$$

Lemma 3.5 is a version of the Ottaviani inequality, see Dudley (1989, p. 251) or Hoffmann-Jørgensen (1994, p. 472). In the form (6) this inequality can be found in Etemadi (1985) with 4 instead of 3 (twice). The proof of Lemma 3.5 repeats those from the references above and is therefore omitted.

Lemma 3.6. *Let the conditions of Theorem* 2.1 *be satisfied and assume that the vectors* X_k, $k = 1, \ldots, 2^N$, *are constructed by the dyadic procedure described in* (2.22)–(2.36). *Then there exist absolute positive constants* $c_{13}, \ldots, c_{17}$ *such that*

a) *If* $\tau\, d^{3/2}/2^{N/2} \leqslant c_9$, *then*

$$\left|U_{N,0} - V_{N,0}\right| \leqslant c_{13}\,d^{3/2}\,\tau\left(1 + 2^{-N}\left|U_{N,0}\right|^2\right) \tag{7}$$

provided that $\left|U_{N,0}\right| \leqslant \dfrac{c_{14}\cdot 2^N}{d^{3/2}\,\tau}$;

b) *If* $1 \leqslant n \leqslant N$, $0 \leqslant k < 2^{N-n}$, $\tau\, d^{3/2}/2^{n/2} \leqslant c_{15}$, *then*

$$\left|\widetilde{U}_{n,k} - \widetilde{V}_{n,k}\right| \leqslant c_{16}\,d^{3/2}\,\tau\left(1 + 2^{-n}\left|\mathbf{U}_{n,k}\right|^2\right) \tag{8}$$

provided that $\left|\mathbf{U}_{n,k}\right| \leqslant \dfrac{c_{17}\cdot 2^n}{d^{3/2}\,\tau}$.

In the proof of Lemma 3.6 we need the following auxiliary Lemma 3.7 which is useful for the application of Lemma 3.4 to the conditional distributions involved in the dyadic scheme.

Corollary 2.1. *Under the conditions of Lemma* 2.2

$$\big|\,S_m - T_m\,\big| \leqslant \big|U_{N,0} - V_{N,0}\,\big| + \frac{1}{2}\sum_{n=r+1}^{N} \big|\widetilde{U}_{n,l_{n,m}} - \widetilde{V}_{n,l_{n,m}}\,\big|, \qquad m = 1,\dots,2^N.$$

This statement evidently follows from Lemmas 2.1 and 2.2 and from the relations (9)–(12), (22) and (23).

3. Proof of Theorem 2.1

In the proof of Theorem 2.1 we shall use the following auxiliary Lemmas 3.1–3.4 (Zaitsev 1995, 1996a,b).

Lemma 3.1. *Suppose that* $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $y \in \mathbb{R}^m$, $\alpha \in \mathbb{R}^1$, *and* $\mathbb{M} : \mathbb{R}^d \to \mathbb{R}^m$ *is a linear operator. Let* $\widetilde{\xi} \in \mathbb{R}^k$ *be the vector consisting of a subset of coordinates of the vector* ξ. *Then*

$$\mathcal{L}(\mathbb{M}\xi + y) \in \mathcal{A}_m\big(\,\|\mathbb{M}\|\,\tau\big), \qquad \text{where} \quad \|\mathbb{M}\| = \sup_{\|x\|\leqslant 1} \|\mathbb{M}x\|\,,$$
$$\mathcal{L}(\alpha\,\xi) \in \mathcal{A}_d(|\alpha|\,\tau), \qquad \mathcal{L}(\widetilde{\xi}) \in \mathcal{A}_k(\tau).$$

Lemma 3.2. *Suppose that independent random vectors* $\xi^{(k)}$, $k = 1, 2$, *satisfy the condition* $\mathcal{L}\big(\xi^{(k)}\big) \in \mathcal{A}_{d_k}(\tau)$. *Let* $\xi = \big(\xi^{(1)},\, \xi^{(2)}\big) \in \mathbb{R}^{d_1+d_2}$ *be the vector with the first* d_1 *coordinates coinciding with those of* $\xi^{(1)}$ *and with the last* d_2 *coordinates coinciding with those of* $\xi^{(2)}$. *Then* $\mathcal{L}(\xi) \in \mathcal{A}_{d_1+d_2}(\tau)$.

Lemma 3.3 (Bernstein-type inequality). *Suppose that* $\mathcal{L}(\xi) \in \mathcal{A}_1(\tau)$, $\mathbf{E}\,\xi = 0$ *and* $\mathbf{E}\,\xi^2 = \sigma^2$. *Then*

$$\mathbf{P}\,\big\{\,|\xi| \geqslant x\,\big\} \leqslant 2\,\max\Big\{\,\exp\big(-x^2/4\sigma^2\big),\ \exp\big(-x/4\tau\big)\,\Big\}, \qquad x \geqslant 0.$$

Lemma 3.4. *Let the distribution of a random vector* $\xi \in \mathbb{R}^d$ *with* $\mathbf{E}\,\xi = 0$ *satisfy the condition* $\mathcal{L}(\xi) \in \overline{\mathcal{A}}_d(\tau, 4)$, $\tau \geqslant 0$. *Assume that the variance* $\sigma^2 = \mathbf{E}\,\xi_d^2 > 0$ *of the last coordinate* ξ_d *of the vector* ξ *is the minimal eigenvalue of* $\operatorname{cov}\xi$. *Then there exist absolute positive constants* $c_8, \dots, c_{12}$ *such that the following assertions hold*:

a) *Let* $d \geqslant 2$. *Assume that* ξ_d *is not correlated with previous coordinates* $\xi_1, \dots, \xi_{d-1}$ *of the vector* ξ. *Define* $\mathbb{B} = \operatorname{cov}\overline{\mathbb{P}}_{d-1}\xi$ *and denote by* $F(z\,|\,x)$, $z \in \mathbb{R}^1$, *the r.c.d.f. of* ξ_d *for a given value of* $\overline{\mathbb{P}}_{d-1}\,\xi = x \in \mathbb{R}^{d-1}$. *Let* $\mathcal{L}(\overline{\mathbb{P}}_{d-1}\,\xi) \in \overline{\mathcal{A}}_{d-1}(\tau, 4)$. *Then there exists* $y \in \mathbb{R}^1$ *such that*

$$|y| \leqslant c_8\,\tau\,\Big\|\,\mathbb{B}^{-1/2}\,x\,\Big\|^2 \leqslant c_8\,\tau\,\frac{\|x\|^2}{\sigma^2}\,, \tag{1}$$

and

$$\Phi_\sigma\big(z - \gamma(z)\big) < F(z + y\,|\,x) < \Phi_\sigma\big(z + \gamma(z)\big), \tag{2}$$

Lemma 2.1 (Einmahl (1989)). *The joint distribution of the constructed vectors* $U_{n,k}$ *and* $\mathbf{U}_{n,k}$ *coincides with that of the vectors* $U^*_{n,k}$ *and* $\mathbf{U}^*_{n,k}$. *In particular,* $X_k,\ k = 1, \dots, 2^N,$ *are independent and* $\mathcal{L}(X_k) = \mathcal{L}(\xi_k)$.

Moreover, according to (11) and (12), we have

$$\begin{aligned} &\widetilde{U}_{n,k} = U_{n-1,2k} - U_{n-1,2k+1}, \\ &U_{n,k} = U_{n-1,2k} + U_{n-1,2k+1} = S_{(k+1)\cdot 2^n} - S_{k\cdot 2^n}, \end{aligned} \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N \tag{37}$$

(it is clear that (37) follows from (35)). Furthermore, putting

$$\widetilde{\mathbf{U}}_{n,k} = \big(U_{n-1,2k},\, U_{n-1,2k+1}\big) \in \mathbb{R}^{2d}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{38}$$

we have (see (13) and (15))

$$\mathbf{U}_{n,k} = \mathbb{A}\,\widetilde{\mathbf{U}}_{n,k} \in \mathbb{R}^{2d}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{39}$$

Note that it is not difficult to verify that, according to (16),

$$\|\mathbb{A}\| = \frac{1}{\|\mathbb{A}^{-1}\|} = \|\mathbb{A}^*\| = \frac{1}{\|(\mathbb{A}^*)^{-1}\|} = \sqrt{2}, \tag{40}$$

where the asterisk is used to denote the adjoint operator $\mathbb{A}^*$ for the operator $\mathbb{A}$.

Remark 2.1. The conditions of Theorem 2.1 imply the coincidence of the corresponding first and second moments of the vectors $\mathbf{U} = \big\{U_{n,k},\, \widetilde{\mathbf{U}}_{n,k},\, \mathbf{U}_{n,k}\big\}$ and $\mathbf{V} = \big\{V_{n,k},\, \widetilde{\mathbf{V}}_{n,k},\, \mathbf{V}_{n,k}\big\}$ since the vectors $\mathbf{U}$ can be restored from $X_1, \dots, X_{2^N}$ by the same linear procedure which is used for reconstruction of the vectors $\mathbf{V}$ from $Y_1, \dots, Y_{2^N}$. In particular, $\mathbf{E}\,\mathbf{U} = \mathbf{E}\,\mathbf{V} = 0$.

Lemma 2.2 (Einmahl 1989, Lemma 5, p. 55). *Let* $1 \leqslant m = (2\,s + 1) \cdot 2^r \leqslant 2^N$, *where* s, r *are non-negative integers. Then*

$$S_m = \frac{m}{2^N}\, S_{2^N} + \sum_{n=r+1}^{N} \gamma_n\, \widetilde{U}_{n,l_{n,m}}, \tag{41}$$

where $\gamma_n = \gamma_n(m) \in [0, {}^1\!/_2\,]$ *and the integers* $l_{n,m}$ *are defined by*

$$l_{n,m} \cdot 2^n < m \leqslant \big(l_{n,m} + 1\big) \cdot 2^n. \tag{42}$$

The shortest proof of Lemma 2.2 can be obtained with the help of a geometrical approach due to Massart (1989, p. 275).

Remark 2.2. The inequalities (42) give a formal definition of $l_{n,m}$. For a better understanding of the mechanism of the dyadic scheme, one should remember another characterization of these numbers: $U_{n,l_{n,m}}$ is the sum over the block of 2^n summands which contains X_m, the last summand in the sum S_m.

(here $\bigl(F^{(1)}_{N,0}\bigr)^{-1}(t) = \sup\bigl\{x : F^{(1)}_{N,0}(x) \leqslant t\bigr\}$, $0 < t < 1$, and so on). Taking into account that the distributions of the random vectors $\xi_1, \ldots, \xi_{2^N}$ are absolutely continuous, we see that this formula can be rewritten in a more natural form, cf. Sakhanenko (1984, p. 30–31):

$$\begin{aligned} F^{(1)}_{N,0}\bigl(U^{(1)}_{N,0}\bigr) &= \Phi_{2^{N/2}}\bigl(V^{(1)}_{N,0}\bigr), \\ F^{(j)}_{N,0}\bigl(U^{(j)}_{N,0} \,\big|\, U^{(1)}_{N,0}, \ldots, U^{(j-1)}_{N,0}\bigr) &= \Phi_{2^{N/2}}\bigl(V^{(j)}_{N,0}\bigr), \qquad \text{for} \quad 2 \leqslant j \leqslant d. \end{aligned} \tag{30}$$

Suppose that the random vectors

$$U_{n,k} = \bigl(U^{(1)}_{n,k}, \ldots, U^{(d)}_{n,k}\bigr), \qquad 0 \leqslant k < 2^{N-n}, \tag{31}$$

corresponding to blocks containing each 2^n summands with fixed n, $1 \leqslant n \leqslant N$, are already constructed. Now our aim is to construct the blocks containing each 2^{n-1} summands. To this end we define

$$\mathbf{U}^{(j)}_{n,k} = \mathbb{P}_j U_{n,k} = U^{(j)}_{n,k}, \quad 1 \leqslant j \leqslant d, \tag{32}$$

and, for $d+1 \leqslant j \leqslant 2\,d$,

$$\mathbf{U}^{(j)}_{n,k} = \bigl(\widetilde{F}^{(j)}_{n,k}\bigr)^{-1}\bigl(\Phi_{2^{n/2}}\bigl(\mathbf{V}^{(j)}_{n,k}\bigr) \,\big|\, \mathbf{U}^{(1)}_{n,k}, \ldots, \mathbf{U}^{(j-1)}_{n,k}\bigr). \tag{33}$$

It is clear that (33) can be rewritten in a form similar to (30). Then we put

$$\begin{aligned} \mathbf{U}_{n,k} &= \bigl(\mathbf{U}^{(1)}_{n,k}, \ldots, \mathbf{U}^{(2d)}_{n,k}\bigr) \in \mathbb{R}^{2d}, \\ \mathbf{U}^{j}_{n,k} &= \bigl(\mathbf{U}^{(1)}_{n,k}, \ldots, \mathbf{U}^{(j)}_{n,k}\bigr) = \overline{\mathbb{P}}_j\,\mathbf{U}_{n,k} \in \mathbb{R}^j, \qquad j = 1, \ldots, 2\,d, \\ \widetilde{U}^{(j)}_{n,k} &= \mathbf{U}^{(j+d)}_{n,k}, \qquad j = 1, \ldots, d, \\ \widetilde{U}_{n,k} &= \bigl(\widetilde{U}^{(1)}_{n,k}, \ldots, \widetilde{U}^{(d)}_{n,k}\bigr) \in \mathbb{R}^d \end{aligned} \tag{34}$$

and

$$\begin{aligned} U_{n-1,2k} &= \frac{1}{2}\bigl(U_{n,k} + \widetilde{U}_{n,k}\bigr), \\ U_{n-1,2k+1} &= \frac{1}{2}\bigl(U_{n,k} - \widetilde{U}_{n,k}\bigr). \end{aligned} \tag{35}$$

Thus, we have constructed the random vectors $U_{n-1,k}$, $0 \leqslant k < 2^{N-n+1}$. After N steps we obtain the random vectors $U_{0,k}$, $0 \leqslant k < 2^N$. Now we set

$$X_k = U_{0,k-1}, \qquad S_0 = 0, \quad S_k = \sum_{l=1}^{k} X_l, \qquad 1 \leqslant k \leqslant 2^N. \tag{36}$$

for $0 \leqslant k < 2^{N-n},\ 1 \leqslant n \leqslant N,\ d+1 \leqslant j \leqslant 2d$. Put

$$T_k = \sum_{l=1}^{k} Y_l, \qquad 1 \leqslant k \leqslant 2^N; \tag{22}$$

$$\begin{aligned} V_{m,k} &= \left(V_{m,k}^{(1)}, \dots, V_{m,k}^{(d)}\right) = T_{(k+1)\cdot 2^m} - T_{k\cdot 2^m}, \\ &\qquad\qquad 0 \leqslant k < 2^{N-m}, \quad 0 \leqslant m \leqslant N; \\ \widetilde{\mathbf{V}}_{n,k} &= \left(V_{n-1,2k},\, V_{n-1,2k+1}\right) = \left(\widetilde{\mathbf{V}}_{n,k}^{(1)}, \dots, \widetilde{\mathbf{V}}_{n,k}^{(2d)}\right) \in \mathbb{R}^{2d}, \\ &\qquad\qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N; \end{aligned} \tag{23}$$

and

$$\mathbf{V}_{n,k} = \left(\mathbf{V}_{n,k}^{(1)}, \dots, \mathbf{V}_{n,k}^{(2d)}\right) = \mathbb{A}\widetilde{\mathbf{V}}_{n,k} \in \mathbb{R}^{2d}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{24}$$

Note that, according to the definition of $\mathbb{A}$, we have (see (11)–(16) and (22)–(24))

$$\mathbf{V}_{n,k} = \left(V_{n,k},\, \widetilde{V}_{n,k}\right) \in \mathbb{R}^{2d}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{25}$$

where

$$\begin{aligned} V_{n,k} &= V_{n-1,2k} + V_{n-1,2k+1}, \\ \widetilde{V}_{n,k} &= V_{n-1,2k} - V_{n-1,2k+1}, \end{aligned} \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{26}$$

and

$$V_{N,0} = Y_1 + \dots + Y_{2^N}. \tag{27}$$

Thus, the vectors $V_{m,k}, \widetilde{\mathbf{V}}_{n,k}$ and $\mathbf{V}_{n,k}$ are constructed from the vectors $Y_1, \dots, Y_{2^N}$ by the same linear procedure which was used for constructing the vectors $U_{m,k}^*, \widetilde{\mathbf{U}}_{n,k}^*$ and $\mathbf{U}_{n,k}^*$ from the vectors $\xi_1, \dots, \xi_{2^N}$.

It is obvious that, for fixed n and k,

$$\operatorname{cov} \mathbf{U}_{n,k}^* = \operatorname{cov} \mathbf{V}_{n,k} = 2^n\, \mathbb{I}_{2d} \tag{28}$$

and, hence, the coordinates of the Gaussian vector $\mathbf{V}_{n,k}$ are independent with the same distribution function $\Phi_{2^{n/2}}(\cdot)$ (here and below

$$\Phi_\sigma(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\,\pi}\,\sigma}\, \exp\Big(-\frac{y^2}{2\,\sigma^2}\Big)\, dy, \qquad x \in \mathbb{R}^1, \quad \sigma > 0,$$

is the distribution function of the normal law with mean zero and variance σ^2).

Denote now the new collection of random vectors X_k as follows. At first we define

$$\begin{aligned} U_{N,0}^{(1)} &= \left(F_{N,0}^{(1)}\right)^{-1}\left(\Phi_{2^{N/2}}\left(V_{N,0}^{(1)}\right)\right) \qquad \text{and,} \quad \text{for} \quad 2 \leqslant j \leqslant d, \\ U_{N,0}^{(j)} &= \left(F_{N,0}^{(j)}\right)^{-1}\left(\Phi_{2^{N/2}}\left(V_{N,0}^{(j)}\right) \,\middle|\, U_{N,0}^{(1)}, \dots, U_{N,0}^{(j-1)}\right) \end{aligned} \tag{29}$$

Now we can formulate the main result of the paper.

Theorem 2.1. *Let the conditions described in* (9)–(17) *be satisfied,* $\tau \geqslant 0$ *and* $\mathbf{E}\,\xi_k = 0,\ \operatorname{cov}\xi_k = \mathbb{I}_d,\ k = 1,\dots,2^N$. *Assume that*

$$\mathcal{L}\big(\mathbf{U}_{n,k}^{*j}\big) \in \overline{\mathcal{A}}_j(\tau, 4) \qquad \text{for} \quad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N, \quad d \leqslant j \leqslant 2\,d, \tag{18}$$

and

$$\mathcal{L}\big(\mathbf{U}_{N,0}^{*j}\big) \in \overline{\mathcal{A}}_j(\tau, 4) \qquad \text{for} \quad 1 \leqslant j \leqslant 2\,d. \tag{19}$$

Then there exist absolute positive constants c_5, c_6 *and* c_7 *such that, for* $\tau\, d^{3/2} \leqslant c_5$, *one can construct on a probability space sequences of independent random vectors* $X_1,\dots,X_{2^N}$ *and i.i.d. Gaussian random vectors* $Y_1,\dots,Y_{2^N}$ *so that*

$$\mathcal{L}(X_k) = \mathcal{L}(\xi_k), \quad \mathbf{E}\,Y_k = 0, \quad \operatorname{cov} Y_k = \mathbb{I}_d, \qquad k = 1,\dots,2^N, \tag{20}$$

and

$$\mathbf{E}\,\exp\Big(\frac{c_6\,\Delta(2^N)}{d^{3/2}\,\tau}\Big) \leqslant \exp\big(c_7\,N\,\log^+ d\big), \tag{21}$$

where $\Delta(2^N) = \max\limits_{1\leqslant r\leqslant 2^N} \Big|\sum\limits_{k=1}^{r} X_k - \sum\limits_{k=1}^{r} Y_k\Big|$.

Theorem 2.1 says that the conditions (18) and (19) suffice for Assertion A. However, these conditions require that the number of summands is 2^N. For an arbitrary number of summands, one should consider additional (for simplicity, Gaussian) summands in order to apply Theorem 2.1.

Below we shall prove Theorem 2.1. Suppose that its conditions are satisfied.

At first, we describe a procedure of constructing the random vectors $\{U_{n,k}\}$ with $\mathcal{L}\big(\{U_{n,k}\}\big) = \mathcal{L}\big(\{U_{n,k}^*\}\big)$, provided that the vectors $Y_1,\dots,Y_{2^N}$ are already constructed (then we shall define $X_k = U_{0,k-1},\ k = 1,\dots,2^N$). This procedure is an extension of the KMT (1975, 1976) dyadic scheme to the multivariate case due to Einmahl (1989). For this purpose we shall use the so-called Rosenblatt quantile transformation (see Rosenblatt (1952) and Einmahl (1989)).

Denote by $F_{N,0}^{(1)}(x_1) = \mathbf{P}\big\{\mathbb{P}_1\,U_{N,0}^* < x_1\big\},\ x_1 \in \mathbb{R}^1$, the distribution function of the first coordinate of the vector $U_{N,0}^*$. Introduce the conditional distributions, denoting by $F_{N,0}^{(j)}\big(\,\cdot\,\big|\,x_1,\dots,x_{j-1}\big),\ 2 \leqslant j \leqslant d$, the regular conditional distribution function (r.c.d.f.) of $\mathbb{P}_j\,U_{N,0}^*$, given $\overline{\mathbb{P}}_{j-1}\,U_{N,0}^* = (x_1,\dots,x_{j-1})$. Denote by $\widetilde{F}_{n,k}^{(j)}\big(\,\cdot\,\big|\,x_1,\dots,x_{j-1}\big)$ the r.c.d.f. of $\mathbb{P}_j\,\mathbf{U}_{n,k}^*$, given $\overline{\mathbb{P}}_{j-1}\,\mathbf{U}_{n,k}^* = (x_1,\dots,x_{j-1})$,

The dyadic scheme. Let N be a positive integer and $\{\xi_1, \dots, \xi_{2^N}\}$ a collection of d-dimensional independent random vectors. Denote

$$\widetilde{S}_0 = 0; \qquad \widetilde{S}_k = \sum_{l=1}^{k} \xi_l, \quad 1 \leqslant k \leqslant 2^N; \tag{9}$$

$$U^*_{m,k} = \widetilde{S}_{(k+1)\cdot 2^m} - \widetilde{S}_{k\cdot 2^m}, \qquad 0 \leqslant k < 2^{N-m}, \quad 0 \leqslant m \leqslant N. \tag{10}$$

In particular, $U^*_{0,k} = \xi_{k+1}$, $U^*_{N,0} = \widetilde{S}_{2^N} = \xi_1 + \dots + \xi_{2^N}$. In the sequel we shall use the term *block of summands* for a collection of summands with indices of the form $k \cdot 2^m + 1, \dots, (k+1) \cdot 2^m$, where $0 \leqslant k < 2^{N-m}$, $0 \leqslant m \leqslant N$. Thus, $U^*_{m,k}$ is the sum over a block containing 2^m summands. Put

$$\widetilde{U}^*_{n,k} = U^*_{n-1,2k} - U^*_{n-1,2k+1}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{11}$$

Note that

$$U^*_{n-1,2k} + U^*_{n-1,2k+1} = U^*_{n,k}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{12}$$

Introduce the vectors

$$\widetilde{\mathbf{U}}^*_{n,k} = \left(U^*_{n-1,2k},\, U^*_{n-1,2k+1}\right) \in \mathbb{R}^{2d}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{13}$$

with the first d coordinates coinciding with those of the vectors $U^*_{n-1,2k}$ and with the last d coordinates coinciding with those of the vectors $U^*_{n-1,2k+1}$. Similarly, denote

$$\mathbf{U}^*_{n,k} = \left(U^*_{n,k},\, \widetilde{U}^*_{n,k}\right) \in \mathbb{R}^{2d}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{14}$$

Introduce now the coordinate mappings (in other words, projectors) $\mathbb{P}_i : \mathbb{R}^s \to \mathbb{R}^1$ and $\overline{\mathbb{P}}_j : \mathbb{R}^s \to \mathbb{R}^j$, $i, j = 1, \dots, s$, by the relations $\mathbb{P}_i x = x_i$, $\overline{\mathbb{P}}_j x = (x_1, \dots, x_j)$, where $x = (x_1, \dots, x_s) \in \mathbb{R}^s$ (we shall use this notation for $s = d$ or $s = 2d$).

It is easy to see that, according to (11)–(14),

$$\mathbf{U}^*_{n,k} = \mathbb{A}\,\widetilde{\mathbf{U}}^*_{n,k} \in \mathbb{R}^{2d}, \qquad 0 \leqslant k < 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{15}$$

where $\mathbb{A} : \mathbb{R}^{2d} \mapsto \mathbb{R}^{2d}$ is a linear operator defined, for $x = (x_1, \dots, x_{2d}) \in \mathbb{R}^{2d}$, as follows:

$$\begin{aligned} &\mathbb{P}_j\,\mathbb{A}\,x = x_j + x_{d+j}, \qquad && j = 1, \dots, d, \\ &\mathbb{P}_j\,\mathbb{A}\,x = x_j - x_{d+j}, \qquad && j = d+1, \dots, 2d. \end{aligned} \tag{16}$$

Denote

$$\begin{aligned} &\mathbf{U}^{*(j)}_{n,k} = \mathbb{P}_j\,\mathbf{U}^*_{n,k}, \\ &\mathbf{U}^{*j}_{n,k} = \left(\mathbf{U}^{*(1)}_{n,k}, \dots, \mathbf{U}^{*(j)}_{n,k}\right) = \overline{\mathbb{P}}_j\,\mathbf{U}^*_{n,k} \in \mathbb{R}^j, \end{aligned} \qquad j = 1, \dots, 2d. \tag{17}$$

2. The main result

Let $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $\|h\|\,\tau < 1$, $h \in \mathbb{R}^d$. The conjugate distribution $\overline{F} = \overline{F}(h)$ is defined by

$$\overline{F}\{dx\} = \left(\mathbf{E}\, e^{\langle h,\xi\rangle}\right)^{-1} e^{\langle h,x\rangle} F\{dx\}. \tag{1}$$

Sometimes we shall write $F_h = \overline{F}(h)$. It is clear that $\overline{F}(0) = F$. Denote by $\overline{\xi}(h)$ a random vector with $\mathcal{L}\big(\overline{\xi}(h)\big) = \overline{F}(h)$. From (1) it follows that

$$\mathbf{E}\, f\big(\overline{\xi}(h)\big) = \left(\mathbf{E}\, e^{\langle h,\xi\rangle}\right)^{-1} \mathbf{E}\, f(\xi)\, e^{\langle h,\xi\rangle}, \tag{2}$$

provided that $\mathbf{E}\,\big|\, f(\xi)\, e^{\langle h,\xi\rangle}\,\big| < \infty$. It is easy to see that

$$\text{if} \quad U_1, U_2 \in \mathcal{A}_d(\tau), \quad U = U_1\, U_2, \quad \text{then} \quad \overline{U}(h) = \overline{U}_1(h)\, \overline{U}_2(h). \tag{3}$$

Below we shall also use the following subclasses of $\mathcal{A}_d(\tau)$ containing distributions satisfying some special smoothness-type restrictions. Let $\tau \geqslant 0$, $\delta > 0$, $\rho > 0$, $h \in \mathbb{R}^d$. Consider the conditions:

$$\int\limits_{\rho\|t\|\tau\, d\geqslant 1} \big|\widehat{F}_h(t)\big|\, dt \leqslant \frac{(2\,\pi)^{d/2}\,\tau\, d^{3/2}}{\sigma\,(\det\mathbb{D})^{1/2}}, \tag{4}$$

$$\int\limits_{\rho\|t\|\tau\, d\geqslant 1} \big|\widehat{F}_h(t)\big|\, dt \leqslant \frac{(2\,\pi)^{d/2}\,\tau^2\, d^2}{\sigma^2\,(\det\mathbb{D})^{1/2}}, \tag{5}$$

$$\int\limits_{\rho\|t\|\tau\, d\geqslant 1} \big|\big\langle t, v\big\rangle\, \widehat{F}_h(t)\big|\, dt \leqslant \frac{(2\,\pi)^{d/2}\,\big\langle \mathbb{D}^{-1}\, v, v\big\rangle^{1/2}}{\delta\,(\det\mathbb{D})^{1/2}}, \qquad \text{for all} \quad v \in \mathbb{R}^d, \tag{6}$$

where $F_h = \overline{F}(h)$ and $\sigma^2 = \sigma^2(F) > 0$ is the minimal eigenvalue of $\mathbb{D} = \operatorname{cov} F$. Denote by $\overline{\mathcal{A}}_d(\tau, \rho)$ (resp. $\mathcal{A}_d^*(\tau, \delta, \rho)$) the class of distributions $F \in \mathcal{A}_d(\tau)$ such that the condition (4) (resp. (5) and (6)) is satisfied for $h \in \mathbb{R}^d$, $\|h\|\,\tau < 1$. It is easy to see that

$$\mathcal{A}_d^*(\tau, \delta, \rho) \subset \overline{\mathcal{A}}_d(\tau, \rho), \qquad \text{if} \quad \frac{\tau\, d^{1/2}}{\sigma} \leqslant 1. \tag{7}$$

In this paper the class $\overline{\mathcal{A}}_d(\tau, \rho)$ plays the role of the class $\mathcal{A}_d^*(\tau, \delta, \rho)$ which was used by Zaitsev (1995, 1996b), see also Sakhanenko (1984, inequality (49), p. 9) or Einmahl (1989, inequality (1.5)). Note that (2) implies

$$\widehat{F}_h(t) = \mathbf{E}\, e^{\langle it,\, \overline{\xi}(h)\rangle} = \left(\mathbf{E}\, e^{\langle h,\, \xi\rangle}\right)^{-1} \mathbf{E}\, e^{\langle h+it,\, \xi\rangle}. \tag{8}$$

where m is a positive integer. Assume that the distributions $\mathcal{L}(\xi_k)$ can be represented in the form

$$\mathcal{L}(\xi_k) = L^{(k)}\, P, \qquad k = 1, \dots, n, \tag{11}$$

where

$$L^{(k)} \in \mathcal{A}_d(\tau) \quad \text{and} \quad P = \mathcal{L}\Big(\big(\zeta_1 + \dots + \zeta_m\big)/\sqrt{m}\Big). \tag{12}$$

Then there exist a positive b_4 depending on H only and such that $m \geqslant b_4$ implies Assertion A.

Remark 1.1. If all the distributions $L^{(k)}$ are concentrated at zero, then the statement of Theorem 1.3 can be derived from the main results of KMT (1975, 1976) (for $d = 1$) and of Zaitsev (1995, 1996b) (for $d \geqslant 1$).

A consequence of Assertion A is given in Theorem 1.4 below.

Theorem 1.4. *Assume that ξ, $\xi_1, \xi_2, \dots,$ are i.i.d. with a common distribution $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$. Let Assertion A be satisfied for $\xi_1, \dots, \xi_n$ for all n with some c_1, c_2 and c_3 independent of n. Suppose that $\tau\, d^{3/2} \leqslant c_1$. Then there exist a construction such that*

$$\mathbf{P}\left\{ \limsup_{n\to\infty} \frac{1}{\log n} \Big| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \Big| \leqslant c_4\, \tau\, d^{3/2} \log^+ d \right\} = 1 \tag{13}$$

with some absolute constant $c_4 = c_4(c_2, c_3)$.

From a result of Bártfai (1966) it follows that the rate $O(\log n)$ in (13) is the best possible if $\mathcal{L}(\xi)$ is non-Gaussian. In the case of distributions with finite exponential moments this rate was established by Zaitsev (1995, 1996b, Corollary 1). Theorems 1.1–1.3 and 2.1 provide examples of smooth distributions which are close to Gaussian ones and for which the constants corresponding to this rate are arbitrarily small. The existence of such examples has been already mentioned in the one-dimensional case by Major (1978, p. 498).

The paper is organized as follows. In Section 2 we formulate Theorem 2.1. To this end we define at first a class of distributions $\overline{\mathcal{A}}_d(\tau, \rho)$ used in Theorem 2.1. The definition of this class is given in terms of smoothness conditions on the so-called conjugate distributions. Then we describe a multidimensional version of the KMT dyadic scheme, cf. Einmahl (1989). We prove Theorem 2.1 in Section 3. Section 4 is devoted to the proofs of Theorems 1.1–1.4.

Acknowledgment. The authors would like to thank V. Bentkus for very useful discussions.

Note that in Assertion A we do not require that the distributions $\mathcal{L}(\xi_k)$ are identical but we assume that they have the same covariance operators, cf. Einmahl (1989) and Zaitsev (1995, 1996b).

According to (1) and (2), the condition $\mathcal{L}(\xi_k) \in \mathcal{A}_d(\tau)$ with small τ means that $\mathcal{L}(\xi_k)$ are close to the corresponding Gaussian laws. It is easy to see that Assertion A becomes stronger for small τ (see as well Theorem 1.4 below). Passing to the limit as $\tau \to 0$, we obtain a spectrum of statements with the trivial limiting case: if $\tau = 0$ (and, hence, $\mathcal{L}(\xi_k)$ are Gaussian) we can take $X_k = Y_k$ and $\Delta(n) = 0$.

We show that *Assertion* A *is valid under some additional smoothness-type restrictions on* $\mathcal{L}(\xi_k)$. The question about the necessity of these conditions remains open. The case $\tau \geqslant 1$ considered by Zaitsev (1995, 1996b, Theorem 1) does not need conditions of such kind. The formulation of our main result—Theorem 2.1—includes some additional notation. In order to show that the conditions of Theorem 2.1 can be verified in some concrete simple situations, we consider at first three particular applications—Theorems 1.1, 1.2 and 1.3.

Theorem 1.1. *Assume that the distributions* $\mathcal{L}(\xi_k) \in \mathcal{A}_d(\tau)$ *can be represented in the form*

$$\mathcal{L}(\xi_k) = H_k \, G, \qquad k = 1, \ldots, n,$$

where G *is a Gaussian distribution with* $\operatorname{cov} G = b^2 \, \mathbb{I}_d$ *and* $b^2 \geqslant 2^{10} \, \tau^2 \, d^3 \, \log^+ \frac{1}{\tau}$. *Then Assertion* A *is valid.*

The following example deals with a non-convolution family of distributions approximating a Gaussian distribution for small τ.

Theorem 1.2. *Let* η *be a random vector with an absolutely continuous distribution and density*

$$p_\tau(x) = \frac{\bigl(4 + \tau^2 \, \|x\|^2\bigr) \, \exp\bigl(-\|x\|^2/2\bigr)}{(2\,\pi)^{d/2} \, (4 + \tau^2 \, d)}, \qquad x \in \mathbb{R}^d. \tag{6}$$

Assume that $\mathcal{L}(\xi_k) = \mathcal{L}\bigl(\eta/\gamma\bigr)$, $k = 1, \ldots, n$, *where*

$$\gamma^2 = \frac{\bigl(4 + \tau^2 \, (d+2)\bigr)}{(4 + \tau^2 \, d)}, \qquad \gamma > 0. \tag{7}$$

Then Assertion A *is valid.*

Theorem 1.3. *Assume that a random vector* ζ *satisfies the relations*

$$\mathbf{E}\,\zeta = 0, \qquad \mathbf{P}\bigl\{\,\|\zeta\| \leqslant b_1\,\bigr\} = 1, \qquad \mathcal{L}(\zeta) \in \mathcal{A}_d(b_2) \tag{8}$$

and admits a differentiable density $p(\cdot)$ *such that*

$$\sup_{x \in \mathbb{R}^d} \bigl|\, d_u \, p(x) \bigr| \leqslant b_3 \, \|u\|, \qquad \text{for all} \quad u \in \mathbb{R}^d, \tag{9}$$

with some positive b_1, b_2 *and* b_3. *Let* $\zeta_1, \zeta_2, \ldots$ *be independent copies of* ζ. *Write*

$$\tau = b_2 \, m^{-1/2}, \tag{10}$$

leads to exponential estimates for the tails of distributions (see, e.g., Lemma 3.3 below). On the other hand, if $\mathbf{E}\, e^{\langle h,\xi\rangle} < \infty$, for $h \in A \subset \mathbb{R}^d$, where A is a neighborhood of zero, then $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau(F))$ with some $\tau(F)$ depending on F only.

Throughout we assume that $\tau \geqslant 0$ and $\xi_1, \xi_2, \ldots$ are random vectors with given distributions $\mathcal{L}(\xi_k) \in \mathcal{A}_d(\tau)$ such that $\mathbf{E}\,\xi_k = 0$, $\operatorname{cov}\xi_k = \mathbb{I}_d$, $k = 1, 2, \ldots$. The problem is to construct, for a given n, $1 \leqslant n \leqslant \infty$, on a probability space a sequence of independent random vectors $X_1, \ldots, X_n$ and a sequence of i.i.d. Gaussian random vectors $Y_1, \ldots, Y_n$ with $\mathcal{L}(X_k) = \mathcal{L}(\xi_k)$, $\mathbf{E}\, Y_k = 0$, $\operatorname{cov} Y_k = \mathbb{I}_d$, $k = 1, \ldots, n$, such that, with large probability,

$$\Delta(n) = \max_{1 \leqslant r \leqslant n} \Big| \sum_{k=1}^{r} X_k - \sum_{k=1}^{r} Y_k \Big| \tag{3}$$

is as small as possible.

The aim of the paper is to provide sufficient conditions for the following Assertion A:

Assertion A. *There exist absolute positive constants c_1, c_2 and c_3 such that, for $\tau\, d^{3/2} \leqslant c_1$, there exists a construction such that*

$$\mathbf{E}\, \exp\Big(\frac{c_2\, \Delta(n)}{d^{3/2}\, \tau} \Big) \leqslant \exp\big(c_3\, \log^+ d\, \log^+ n \big). \tag{4}$$

Using the exponential Chebyshev inequality, we see that (4) implies

$$\mathbf{P}\Big\{ c_2\, \Delta(n) \geqslant \tau\, d^{3/2} \big(c_3\, \log^+ d\, \log^+ n + x \big) \Big\} \leqslant e^{-x}, \qquad x \geqslant 0. \tag{5}$$

Therefore, Assertion A can be considered as a generalization of the classical result of KMT (1975, 1976). Assertion A provides a supplement to an improvement of a multidimensional KMT-type result of Einmahl (1989) presented by Zaitsev (1995, 1996b) which differs from Assertion A by the restriction $\tau \geqslant 1$ and by another explicit power-type dependence of the constants on the dimension d. In a particular case, when $d = 1$ and all summands have a common variance, the result of Zaitsev is equivalent to the main result of Sakhanenko (1984), who extended the KMT construction to the case of non-identically distributed summands and stated the dependence of constants on the distributions of the summands belonging to a subclass of $\mathcal{A}_1(\tau)$.

The main difference between Assertion A and the aforementioned results consists in the fact that in Assertion A we consider "small" τ, $0 \leqslant \tau \leqslant c_1\, d^{-3/2}$. In previous results the constants were separated from zero by quantities which are larger than some absolute constants. In KMT (1975, 1976) the dependence of the constants on the distributions is not specified. From the conditions (1) and (4) in Sakhanenko (1984), Section 1, it follows that $\operatorname{Var}\xi_k \leqslant \lambda^{-2}$ (λ^{-1} plays in Sakhanenko's paper the role of τ) and, if $\operatorname{Var}\xi_k = 1$, then $\lambda^{-1} \geqslant 1$. This corresponds to the restrictions $\alpha^{-1} \geqslant 2$ in Einmahl (1989, conditions (3.6) and (4.3)) and $\tau \geqslant 1$ in Zaitsev (1995, 1996b, Theorem 1).

the representation $z = (z_1, \dots, z_d) = z_1 e_1 + \cdots + z_d e_d$, where $z_j \in \mathbb{R}^1$ (resp. $\mathbb{C}^1$) and the e_j are the standard orthonormal vectors. The scalar product is denoted by $\langle x, y \rangle = x_1 \overline{y}_1 + \cdots + x_d \overline{y}_d$. We shall use the Euclidean norm $\|z\| = \langle z, z \rangle^{1/2}$ and the maximum norm $|z| = \max\limits_{1 \leqslant j \leqslant k} |z_j|$. The symbols $c, c_1, c_2, \dots$ will be used for absolute positive constants. The letter c may denote different constants when we do not need to fix their numerical values. The ends of proofs will be denoted by $\square$. The formulas are enumerated in each section separately. When we refer to a formula from another section, we indicate in addition the number of this section. For example, (2.3) is the formula (3) from Section 2.

Let us consider the definition and some useful properties of classes of distributions $\mathcal{A}_d(\tau) \subset \mathfrak{F}_d$, $\tau \geqslant 0$, introduced in Zaitsev (1986), see as well Zaitsev (1995, 1996a,b). The class $\mathcal{A}_d(\tau)$ (with a fixed $\tau \geqslant 0$) consists of distributions $F \in \mathfrak{F}_d$ for which the function

$$\varphi(z) = \varphi(F, z) = \log \int_{\mathbb{R}^d} e^{\langle z, x \rangle} F\{dx\} \qquad (\varphi(0) = 0)$$

is defined and analytic for $\|z\| \, \tau < 1$, $z \in \mathbb{C}^d$, and

$$\bigl| d_u d_v^2 \, \varphi(z) \bigr| \leqslant \|u\| \, \tau \, \langle \mathbb{D} v, v \rangle \qquad \text{for all } \ u, v \in \mathbb{R}^d \ \text{ and } \ \|z\| \, \tau < 1,$$

where $\mathbb{D} = \operatorname{cov} F$, and the derivative $d_u \varphi$ is given by

$$d_u \varphi(z) = \lim_{\beta \to 0} \frac{\varphi(z + \beta u) - \varphi(z)}{\beta} \, .$$

It is evident that $\tau_1 < \tau_2$ implies $\mathcal{A}_d(\tau_1) \subset \mathcal{A}_d(\tau_2)$. Furthermore, the class $\mathcal{A}_d(\tau)$ is closed with respect to convolution: if $F_1, F_2 \in \mathcal{A}_d(\tau)$, then $F_1 F_2 \in \mathcal{A}_d(\tau)$. The class $\mathcal{A}_d(0)$ coincides with the class of all Gaussian distributions in $\mathbb{R}^d$. The following inequality can be considered as an estimate of the stability of this characterization: if $F \in \mathcal{A}_d(\tau)$, $\tau > 0$, then

$$\pi\bigl(F, \, \Phi(F)\bigr) \leqslant c d^2 \tau \, \log^+ \tau, \tag{1}$$

where $\pi(\cdot, \cdot)$ is the Prokhorov distance and $\Phi(F)$ denotes the Gaussian distribution whose mean and covariance operator are the same as those of F. Moreover, for all $X \in \mathfrak{B}_d$ and all $\lambda > 0$, we have

$$\begin{aligned} F\{X\} &\leqslant \Phi(F)\{X^\lambda\} + c d^2 \exp\Bigl(- \frac{\lambda}{c d^2 \, \tau} \Bigr), \\ \Phi(F)\{X\} &\leqslant F\{X^\lambda\} + c d^2 \exp\Bigl(- \frac{\lambda}{c d^2 \, \tau} \Bigr), \end{aligned} \tag{2}$$

where $X^\lambda = \Bigl\{ y \in \mathbb{R}^d : \inf\limits_{x \in X} \|x - y\| < \lambda \Bigr\}$ is the λ-neighborhood of the set X, see Zaitsev (1986).

The classes $\mathcal{A}_d(\tau)$ are closely connected with other natural classes of multidimensional distributions. In particular, by the definition of $\mathcal{A}_d(\tau)$, a distribution $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$ has finite exponential moments $\mathbf{E} \, e^{\langle h, \xi \rangle}$, for $\|h\| \, \tau < 1$. This

MULTIDIMENSIONAL HUNGARIAN CONSTRUCTION FOR VECTORS WITH ALMOST GAUSSIAN SMOOTH DISTRIBUTIONS

F. GÖTZE[1] A. YU. ZAITSEV[2]

University of Bielefeld[1]
St. Petersburg Branch of Steklov Mathematical Institute[2]

June 1997

ABSTRACT. A multidimensional version of the results of Komlós, Major and Tusnády for sums of independent random vectors with finite exponential moments is obtained in the particular case where the summands have smooth distributions which are close to Gaussian ones. The bounds obtained reflect this closeness. Furthermore, the results provide sufficient conditions for the existence of i.i.d. vectors $X_1, X_2, \dots$ with given distributions and corresponding i.i.d. Gaussian vectors $Y_1, Y_2, \dots$ such that, for given small ε,

$$\mathbf{P}\left\{\limsup_{n\to\infty}\frac{1}{\log n}\Big|\sum_{j=1}^{n}X_j-\sum_{j=1}^{n}Y_j\Big|\leqslant\varepsilon\right\}=1.$$

1. Introduction

The paper is devoted to an improvement of a multidimensional version of strong approximation results of Komlós, Major and Tusnády (KMT) for sums of independent random vectors with finite exponential moments and with smooth distributions which are close to Gaussian ones.

Let $\mathfrak{F}_d$ be the set of all d-dimensional probability distributions defined on the σ-algebra $\mathfrak{B}_d$ of Borel subsets of $\mathbb{R}^d$. By $\widehat{F}(t)$, $t\in\mathbb{R}^d$, we denote the characteristic function of a distribution $F\in\mathfrak{F}_d$. The product of measures is understood as their convolution, that is, $F\,G=F*G$. The distribution and the corresponding covariance operator of a random vector ξ will be denoted by $\mathcal{L}(\xi)$ and $\operatorname{cov}\xi$ (or $\operatorname{cov}F$, if $F=\mathcal{L}(\xi)$). The symbol $\mathbb{I}_d$ will be used for the identity operator in $\mathbb{R}^d$. For $b>0$ we denote $\log^+ b=\max\left\{1,\log b\right\}$. Writing $z\in\mathbb{R}^d$ (resp. $\mathbb{C}^d$), we shall use

1991 *Mathematics Subject Classification.* Primary 62F17; secondary 60F15.

Key words and phrases. Multidimensional invariance principle, strong approximation, sums of independent random vectors, Central Limit Theorem.

[1]Research supported by the SFB 343 in Bielefeld by grant INTAS-RFBR 95-0099 and by grant RFBR-DFG 96-01-00096-ge.

[2]Research supported by by the Russian Foundation of Basic Research, grant 96-01-00672, by grant INTAS-RFBR 95-0099, by grant RFBR-DFG 96-01-00096-ge and by the SFB 343 in Bielefeld.

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX